The Dissertation Committee for Lalita Maria Calabria certifies that this is the approved version of the following dissertation:

The Isolation and Characterization of Triterpene Saponins from *Silphium* and the Chemosystematic and Biological Significance of Saponins in the Asteraceae

Committee:

Tom J. Mabry, Supervisor

Stanley Roux, Co-Supervisor

Joe Karchesy

Alan Lloyd

Ben Shoulders

The Isolation and Characterization of Triterpene Saponins from *Silphium* and the Chemosystematic and Biological Significance of Saponins in the Asteraceae

by

Lalita Maria Calabria, B.S.

Dissertation

Presented to the Faculty of Graduate School of

the University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

May, 2008

This dissertation is dedicated to my Nanny and Pop-pop,

whose hard work and sacrifice paved

the way for my future.

Your unconditional love and support

are the glue that holds our family together.

"To see a world in a grain of sand,
And a heaven in a wild flower,
Hold infinity in the palm of your hand,
An eternity in an hour"
-William Blake

Acknowledgments

This dissertation represents one of the greatest adventures of my life. Many people have inspired me to study the remarkable properties of plants and I am grateful to all of them. I want to begin by acknowledging Dr. Peter J. Pessiki, who first introduced me to the mysteries of natural product chemistry. Thank you for challenging me to pursue graduate school. I can finally answer your question: Yes, I have what it takes and yes, I would like fries with that.

I would like to express my deepest appreciation to my advisor, Dr. Tom J. Mabry who exemplifies the personal and professional integrity to which I aspire. I am very grateful for his constant support over the last five years and during the completion of this dissertation. His patience in reading draft after draft of every manuscript, proposal and idea I wrote continues to amaze me. No one should be subjected to the torture of reading my early attempts at technical writing, and thanks to Dr. Mabry, no one will. I am grateful that Dr. Mabry gave me the freedom to pursue my own research ideas and reigned in my imagination when it got the better of me. He has taught me, through example, how to be successful in academia and in life. I am forever grateful for his mentorship and I am truly honored to have been his last student.

I wish to thank the members of my committee, Joe Karchesy, Alan Lloyd, Stan Roux and Ben Shoulders for their guidance since my beginning days as a graduate student. Their examples, both as researchers and teachers, continue to serve as guidelines for my academic career. I am especially grateful to my co-supervisor, Dr. Roux for opening up his lab space for me to finish writing my dissertation.

It is my pleasure to thank my collaborators, Su Dharmawardane, Vicente Emerenciano, Irenuesz Kapusta, Venkat Krishnamachari, Wieslaw Oleszek, Paul Paré, Peter Pessiki, Sonia Piacente, Marcus Scotti and Frances Segarra. Much of the research

described in this dissertation is jointly theirs. I owe a special gratitude to Dr. Jeffrey Williams, Dr. Jennifer Clevinger, Dr. Malgorzata Wojcińska, Dr. Kenneth Roberston, and Don Gardner, who collected or grew most of the plant material used in my studies.

I am grateful for the talented and dedicated visiting scientists that taught me how to "work smart" in the lab. In particular, I wish to thank the late Prof. Ahmed A. Ahmed who patiently guided me through each step of building my first column for isolating saponins; to Dr. Nabil El-Sayed for the long hours spent interpreting NMR spectra. Thank you for sharing those eureka moments with me. And to Dr. Malgorzata Wojcińska for pushing me to come in earlier, stay up later and get more done in the lab. Without you, I would still be sitting in front of five columns collecting fractions.

There are a few professors who were not on my committee but went above and beyond to help me through this journey. Dr. Jerry Brand graciously allowed me to work in his lab when there was no alternative space available. Dr. José Panero provided me with a passionate education on the Asteraceae whenever I needed some inspiration and, on extremely short notice, sat in as a member of my preliminary exam committee. I am especially grateful to Dr. Beryl Simpson for her editorial comments on both manuscripts and chapters and for many entertaining and enlightening conversations.

I am forever indebted to the several dozen, bright and talented undergraduate research assistants who kept the lab running twelve hours a day, six days a week, for nearly two straight years. Without their hard work and dedication, I would not have been able to complete the isolations and structural elucidations necessary to characterize saponins from all *Silphium* species. It is extremely rewarding for me to see many of these students going on to graduate school in pharmacy, medical school and other successful research careers.

I owe special thanks to Dr. Mehdi Moini and his staff in the Mass Spec facility at UT-Austin for their assistance in obtaining important mass spectral data. I am particularly grateful to Dr. Klaus Linse (Protein Core Facility) for his friendship and for patiently teaching me the "ins and outs" of HPLC and MS instrumentation. The many hours you spent fixing our HPLC and running our LC-MS samples are very much appreciated.

I am also grateful to the staff and technicians at the NMR facility. In particular, Steve Sorey and Jim Wallin, who afforded NMR spectra of consistently high quality, despite the questionable purity of some of the samples I submitted. Finally, I would like to thank Dr. Ben Shoulders for always having his door open and for his patience and assistance with interpreting my NMR spectra.

I would like to thank the Plant Biology Graduate Coordinator, Tamra Rogers, and the entire MCDB staff, including Sasha Bennet, Dulce Marquez, Maureen Meko, Francisco Valladares and Kendrick Weld for their friendly assistance, making our department a great place to work.

I am blessed with the love and support of many loyal friends, who have humored me through years of "nerding out" with plants and neglecting them. Now that I'm done writing my dissertation, I promise that I will call/visit more. I am especially grateful to all of my peers in the Plant Biology graduate program, who were always available for inspiration or distraction. In particular, I want to thank Schonna Manning, Mary Guisinger-Bellian, Aimee Talarski and "Scoop Daddy" for our "bitch lunches" and all the important gossiping that was accomplished.

The hard-work and dedication required for this endeavor was instilled in me by my parents, Charlie and Cathy, who taught me, by example, how to be a success. Thank you for always encouraging my natural curiosity and creativity; these qualities have been essential to my PhD studies and have made my life in general, SO much fun. I am

especially grateful to my mom for her unconditional, all encompassing love and support and for making me follow through with everything, always. And to my step-dad Bob, for all of his joking and poking that kept me from taking myself too seriously. Thank you both for always believing in me. And to my daddy, who has been a pillar of strength and integrity in my life. Your love and support means everything to me. I will always be your little girl in the donut shop with the powdered sugar all over my face- except now you'll have to call me "Dr."! And, finally to my brother Vinnie: thank you for loving me more than anyone in the entire world.

I am blessed to be a part of another loving and supportive family: The Pinchot's. My father-in-law Gifford was instrumental in helping me to choose my next step after graduating from college. Thank you for encouraging me to go after my biggest dreams and whole-heartedly believing that I could achieve them. To my mother-in-law, Libba, thank you for welcoming me into your family and sharing your strength and your vulnerability with me. In some of my deepest moments of self-doubt, you reminded me that I can do anything and I am grateful for your continual support throughout my doctoral studies. To my brother and sister-in-law, Alex and Marianna: thank you for taking care of me and for all the joy and laughter you have bestowed on me.

I am indebted to my husband Marco for his enduring patience, love and encouragement throughout my PhD studies. I cannot imagine a more supportive husband. Thank you for moving to Texas, for cooking and cleaning, for tolerating the long hours and short fuses, and for making me laugh every single day. Your love and support have made everything in my life possible and I can hardly wait to begin our next big adventure!

Financial support for this work was provided in part by the Robert A. Welch Foundation grant (F-130) to Tom J. Mabry and a University Continuing Fellowship (2007-2008).

The Isolation and Characterization of Triterpene Saponins from *Silphium* and the Chemosystematic and Biological Significance of Saponins in the Asteraceae

Publication No._____

Lalita Maria Calabria, Ph.D.

The University of Texas at Austin, 2008

Supervisors: Tom J. Mabry, Stanley Roux

My dissertation studies involved the isolation and characterization of triterpene saponins from eleven species of *Silphium* L. (Asteraceae). Saponins are an economically important group of natural products with a wide range of biological activities and a broad distribution in the plant kingdom. Typically, saponins consist of one or more sugar moieties attached to either a steroid or a triterpene aglycone. In the course of these investigations, nine new triterpene saponins were identified. The compounds were identified as 3β,6β,16β-trihydroxyolean-12-en-23-al-3-*O*-β-glucopyranosyl-16-*O*-β-glucopyranoside (**1**), urs-12-ene-3β,6β,16β-triol-3-*O*-β-galactopyranosyl-(1 → 2)-β-glucopyranoside (**2**), 3β,6β,16β-trihydroxyolean-12-en-23-oic acid-3-*O*-β-glucopyranosyl-16-*O*-β-glucopyranoside (**3**), urs-12-ene-3β,6β,16β,21β-tetraol-3-*O*-β-glucopyranoside (**4**), olean-12-ene-3β,6β,16β,21β-tetraol-3-*O*-β-glucopyranoside (**5**),

olean-12-ene-3β,6β,16β,21β,23-pentaol-3-*O*-β-glucopyranosyl-16-*O*-β-glucopyranoside (**6**), olean-12-ene-3β,6β,16β-triol-3-*O*-β-glucopyranosyl-16-*O*-α-arabinopyranosyl-(1 → 2)-β-glucopyranoside (**7**), olean-12-ene-3β,6β,16β,23-tetraol-3-*O*-β-glucopyranosyl-16-*O*-α-arabinopyranosyl-(1 → 2)-β-glucopyranoside (**8**) and 3β,6β,16β,21β-tetrahydroxyolean-12-en-23-al-3-*O*-β-glucopyranoside (**9**). Moreover, the cytotoxic activities of the isolated compounds were tested against human breast cancer cell line MDA-MB-231. Results showed that compound **2** decreased breast cancer cell proliferation in a statistically significant manner at 25 μg/ml. Several other known saponins were also isolated and identified. The structures of these compounds were established by chemical methods and spectral analyses including ^{1}H-NMR, ^{13}C NMR, HMBC, NOESY DEPT and TOSCY experiments as well as high resolution and low resolution ESI-MS analyses. In addition to the full structural characterization of the above-mentioned compounds, a comprehensive genus-wide HPLC/ESI-MS analysis of *Silphium* leaf extracts revealed the presence of over 90 additional saponins in the *Silphium* genus and in the closely related *Lindheimera texana*. To complement my chemical studies of *Silphium*, I also present a phylogenetic analysis of the entire Asteraceae based on phytochemical data and a detailed overview of the chemistry of the Asteraceae using a large-scale chemical database to view the phytochemical data within the framework of the most recent molecular phylogeny for the Asteraceae. Finally, I comment on the results of my studies in the context of the chemosystematic and medicinal significance of triterpene saponins from the Asteraceae.

Table of Contents

List of Tables……xiv

List of Figures…….…….....……xv

Abbreviations……xix

Chapter 1: Introduction……1

1.1 Summary of Chapters……1
1.2 Why I Chose to Investigate Triterpene Saponins in *Silphium* for my Dissertation……2
1.3 Research Objectives……...6

Chapter 2: Background……8

A. *Silphium* L. (Asteraceae) ……..……8

2.1 Natural History……...…8
2.2 Taxonomic History and Current Phylogenetic Placement……...…13
2.3 Ethnobotany…….……15
2.4 Chemistry…….…17

2.4.1 Diterpenes…….…18
2.4.2 Sesquiterpenes…….……...…18
2.4.3 Flavonoids……...…….……19
2.4.4 Phenolic acids……..……20
2.4.5 Monoterpenes……20
2.4.6 Triterpenes……22

B. Saponins……..……23

2.5 Historical Overview of Saponins…….24
2.6 Structures and Biosynthesis of Saponins…….……24
2.7 Functions of Saponins in Plants……...…27
2.8 Medicinal Actions and Bioactivities of Saponins……29

2.81 Anti-cancer Activity……...…29
2.8.2 Anti-inflammatory……...……29
2.8.3 Cardiovascular Activity……30
2.8.4 Anti-viral, Anti-microbial and Anti-fungal Activities…….…….30

2.9 Saponins as Chemosystematic Markers……...……31

Chapter 3: A Phylogenetic Analysis of Tribes of the Asteraceae Based on Phytochemical Data……..……33

3.1 Introduction……..……33
3.2 Methods……..…..……41
3.3 Results and Discussion……45
3.4 Conclusions……..……50

Chapter 4: Secondary Chemistry of the Asteraceae……..…54

4.1 Introduction……..…55
4.2 Background and Historical Perspectives…….……55
4.3 Methods……62
4.4 Results and Discussion……...…….64
4.5 Conclusions……..……86

Chapter 5: Triterpene saponins from *Silphium* L. (Asteraceae)…….……...88

5.1 Triterpene saponins from *Silphium radula*…….…….88

5.1.1 Introduction……...…89
5.1.2 Results and Discussion……..…90
5.1.3 Experimental…….…..…106

5.2 The Triterpene Saponin Chemistry of *Silphium terebinthinaceum* and *S. laciniatum*…….……111

5.2.1 Introduction…….……111
5.2.2 Methods……...…113
5.2.3 Results……..…116
5.2.4 Discussion and Future Work……...121

5.3. Isolation and Characterization of Triterpene Saponins from *Silphium integrifolium, S. morhii* and *S. perfoliatum* and Discussion of Preliminary Chemical Analysis of Triterpene Saponins from Five Other *Silphium* Species…….….127

5.3.1 *Silphium integrifolium*……128
5.3.2 *Silphium morhii*…….……..…132
5.3.3 *Silphium perfoliatum*…….……...…135
5.3.4 Experimental..…….……138

Chapter 6: Metabolic Profiling of Triterpene Saponins in the Genus *Silphium* and the Closely Related Species *Lindheimera texana* by Electrospray Ion-Trap Mass Spectrometry……..…142

6.1 Introduction……...….142
6.2 Materials and Methods……...…148
6.3 Results and Discussion……..……151

6.3.1 Profiling of *S. radula*……..……151
6.3.2 Profiling of Saponins from all *Silphium* species and *Lindheimera texana*……..……155
6.3.3 Additional Ions of Interest……..…170
6.3.4 Saponins from Kowalski's (2007) LC-MS Studies of *Silphium* Species……175

6.4 Conclusions……177

Chapter 7: Conclusions……179

Appendix A……...…185

Appendix B……..…231

Appendix C……..……253

Appendix D……..…261

References……283

Vita……..……301

List of Tables

Table 2.1. Summary of the ethnobotanical uses of *Silphium* species by Native American tribes……..……..……16

Table 3.1. Biphyletic classification systems of Carlquist (1976) and Wagenitz (1976) based on morphological characters…….....……..…36

Table 3.2. The taxa analyzed in this study and their respective three-letter acronyms, the number of species in each taxa and the number of species in the database.…43

Table 4.1. Tribes of the Asteraceae analyzed in this study and their respective three-letter acronyms, the number of species in each tribe and the number of species recorded in our chemical database……...……63

Table 4.2. Raw data matrix created for this study showing the number of occurrences of each class of secondary metabolites recorded in our chemical database from each tribe of the Asteraceae…….....……64

Table 5.1. ESI-MS results for the analysis of saponins from roots of *S. laciniatum* (LACI) and *S. terebinthinaceum* (TERE)……...……121

Table 5.2. HPLC/ESI-MS negative mode data showing intraspecific variation of saponins for *S. perfoliatum* from the United States and from Poland……..……136

Table 6.1. LC-MS retention times and mass spectral characteristics for saponins **A-I**, all previously reported from *Silphium radula*……...……152

Table 6.2. Summary table of HPLC/ESI-MS results for *Silphium* species analyzed in this study…….....……160

Table 6.3. Table comparing saponins detected in extracts of *Silphium* species (A) the current LC-MS analysis with (B) Kowalski's (2007) analysis…….....…176

List of Figures

Figure 2.1. Resinous gum exudes from the stem of *S. laciniatum*……..…….……8

Figure 2.2. Longitudinal section of fruiting head of *S. integrifolium* illustrating the mature fruits and phyllaries subtending the individual fruits……..……9

Figure 2.3. The phylogenetic tree for *Silphium* and the related genera *Lindheimera* based on ITS and ETS sequence data……...…….....15

Figure 2.4. An example a saponin typically found in *Silphium perfoliatum*, silphioside B, first isolated by Davidyants et al., 1984a.……..…….....22

Figure 2.5. Examples of steroidal (A) and triterpene (B) saponins…….....….....…25

Figure 2.6. A diagram representing the biosynthetic pathway leading from mevalonic acid to squalene……..……26

Figure 2.7. A diagram representing the biosynthetic pathway leading to the formation of pentacyclic triterpenes…….....……27

Figure 3.1. Diagram according to Cassini (1816) showing the interrelationships of nineteen tribes of the Asteraceae……...…….……35

Figure 3.2. Diagram according Bentham (1873) reducing the number of tribes to thirteen from Cassini's arrangement (1816) for nineteen tribes. ….....……35

Figure 3.3. Bremer's (1987) strict consensus tree based morphological and chemical characters…….....……37

Figure 3.4. Morphologically-based diagram showing tribes of Asteraceae according to Bremer (1996) …….....……38

Figure 3.5. Kim and Jansen's (1995, 1996) strict consensus ndhF tree condensed to tribes…….....……39

Figure 3.6. The 389 skeletal types used in this study fall into 9 chemical classes represented here by typical compounds found in the Asteraceae.....……42

Figure 3.7. Our strict consensus tree (I) of three equally parsimonious trees with 657 steps based on chemical data using the tribal classification of Bremer (1994)..…46

Figure 3.8. Our strict consensus tree (II) of three shortest trees of 642 steps based on chemical data using the tribal assignments of Kim and Jansen (1995)…..…47

Figure 4.1. An example of the nine major chemical classes in our chemical database represented by typical compounds found in the Asteraceae..……...…56

Figure 4.2. Total chemical occurrences for the Asteraceae representing each class of secondary metabolites in our chemical database…….....…….67

Figure 4.3. Distribution of benzofurans viewed at the tribal level, expressed as "presence" or "absence" of chemical data recorded in our chemical database for the Asteraceae…….....69

Figure 4.4. Distribution of alkaloids viewed at the tribal level, expressed as "presence" or "absence" of chemical data recorded in our database for the Asteraceae…..72

Figure 4.5. Total occurrences for each chemical class found in our chemical database for subfamilies of the Asteraceae according to Panero and Funk (2002)…….....74

Figure 4.6. Individual pie charts for subfamilies Mutisioideae, Carduoideae, Cichorioideae and Asteroideae representing the total chemical occurrences expressed as percentages of each chemical class found in our chemical database for the Asteraceae….....76

Figure 4.7. Total occurrences of each chemical class found in our chemical database for tribes of the Asteraceae. ……..……...…78

Figure 4.8. Total occurrences of each chemical class found in our chemical database for tribes of the Asteraceae, excluding Anthemideae.…….....……79

Figure 4.9. Degree of triterpene oxidation present in the subtribe Engelmanniinae and closely related genera.…….....…….....81

Figure 5.1. Structures of polyhydroxylated triterpene saponins (**1-9**) isolated from *Silphium radula*..…….....…….....……91

Figure 5.2. ^{1}H-NMR spectra and structure for compound **1** isolated from *Silphium radula*, highlighting the aldehydic and olefinic proton resonances...…….....….....…93

Figure 5.3. ^{1}H-NMR spectra for compound **1** isolated from *Silphium radula*, showing the expansion of the sugar region.……...……95

Figure 5.4. HSQC experiment for compound **1** isolated from *Silphium radula*……96

Figure 5.5. Initial testing of the effect of 20 μg/ml of compounds **1-9** on growth of MDA-MB-231 human breast cancer cells…….....103

Figure 5.6a and b. Effect of compound **2** on breast cancer cell proliferation……..105

Figure 5.7. Photographs of leaves of A) *S. laciniatum* B) possible hybrid between *S. laciniatum* and *S. terebinthinaceum* and C) *S. terebinthinaceum*….....112

Figure 5.8. The taproots of *S. laciniatum* (right) and *S. terebinthinaceum*……..……119

Figure 5.9. LC chromatograms and ESI-MS spectra of *S. laciniatum* and *S. terebinthinaceum* root extracts……....……120

Figure 5.10. A longitudinal section of stems of *S. laciniatum and S. terebinthinaceum* revealing *Antistrophus* galls ……...……122

Figure 5.11. Structures of triterpene saponins isolated from *Silphium integrifolium*.…...130

Figure 5.12. Effect of Compound **1** on breast cancer cell proliferation……..……132

Figure 5.13. Structure of compound **1** isolated from *S. morhii*, which was first isolated from *Ilex rotunda* and is known as Ilexoside XLVIII……....……134

Figure 6.1. Structures of triterpene saponins isolated from *S. radula*……....……144

Figure 6.2. Structures of triterpene saponins isolated from *Silphium perfoliatum*….....145

Figure 6.3. MS spectra of the $[M-H]^-$ of literature reported saponins, **A** (*m/z* 813), **F** (*m/z* 913) and **I** (*m/z* 781) from *Silphium radula*..…….....…154

Figure 6.4a. Base peak chromatograms obtained by negative ion HPLC/ESI-MS of leaf extracts of (A) *Silphium asteriscus,* (B) *S. brachiatum,* (C) *S. integrifolium* and (D*) S. morhii.*..……156

Figure 6.4b. Base peak chromatograms obtained by negative ion HPLC/ESI-MS of leaf extracts of (E) *Silphium perfoliatum,* (F) *S. radula,* (G) *S. wasiotense* and (H) *S. albiflorum*.…….....……157

Figure 6.4c. Base peak chromatograms obtained by negative ion HPLC/ESI-MS of leaf extracts of (I) *Silphium laciniatum,* (J) *S. terebinthinaceum,* (K) *S. compositum* and (L) *Lindheimera texana*…….....……158

Figure 6.5. Total saponin concentration in crude extracts of *Silphium* and *Lindheimera texana*..…….....……159

Figure 6.6. Mass spectrum for compound **102**, tentatively assigned as saponin **N**, which was present in all species analyzed in this study…….....……161

Figure 6.7. Mass spectrum for compound **39**, tentatively assigned as saponin **J**, detected in *Silphium morhii*…….....…….....……167

Figure 6.8. Mass spectrum for the unknown compound **13**……..…….…….……171

Figure 6.9. Mass spectrum for compound **24**……..……173

Abbreviations

COSY: Homonuclear shift correlation **(CO**rrelated **S**pectroscop**Y).** ^{1}H-^{1}H-COSY identifies which protons are coupled to each other.

DEPT: Distortion less **E**nhancement by **P**olarization **T**ransfer produces separate carbon subspectra for methyl, methylene and methane signals.

ESI-MS: Electro**S**pray **I**onization **M**ass **S**pectrometry

ETS: External **T**ranscribed **S**pacer. Refers to the 18S-26S nrDNA repeat and is considered a good candidate for augmenting or replacing the ITS when the ITS lacks sufficient phylogenetic signal.

HMBC: Heteronuclear **M**ultiple-**B**ond **C**orrelation. Provides means to identify H-C connectivity through two or more bonds.

HPLC/MS: High **P**erformance **L**iquid **C**hromatography **M**ass **S**pectrometry.

HSQC: Heteronuclear **S**ingle-**Q**uantum **C**oherence. Provides direct H-C shift correlations for data related to the connectivity of the proton to the carbon to which it is directly bound.

ITS: Internal **T**ranscribed **S**pacer. Refers to the 18S-26S nuclear ribosomal DNA (nrDNA) repeat used extensively for plant phylogenetic reconstruction and has proven especially useful for elucidating relationships among closely related genera in Asteraceae.

MDA-MB-231: Human breast cancer cell line.

NOESY: Nuclear **O**verhauser **E**nhancement **S**pectroscop**Y**. Used to establish correlations between nuclear spins that are spatially close.

NMR: Nuclear **M**agnetic **R**esonance

TLC: Thin **L**ayer **C**hromatography

TOCSY: TOtal **C**orrelated **S**pectroscop**Y**. Useful for dividing proton signals into groups based on coupling networks.

Chapter 1: Introduction

1.1 Summary of Chapters

This dissertation is organized into 7 major chapters. Chapter 1 is divided into three sections beginning with a summary of each chapter (section 1.1.), followed by an account of how I became interested in the topics covered in my dissertation (section 1.2). The final section (section 1.3) outlines the major research objectives of my doctoral studies.

Chapter 2 provides a review of the literature data pertaining to *Silphium* including the natural history, systematics, ethnobotany and chemistry of the genus. This chapter also introduces saponins, which are glycosides of steroid or triterpene compounds with a wide range of biological activities and a broad distribution in the plant kingdom. A detailed discussion of saponins is presented including structural information, bioactivities and physiological and ecological functions. A review of previous literature on saponins found in *Silphium* is also presented.

Chapters 3 and 4 together present two different approaches for utilizing phytochemical data as a tool for understanding the evolution and diversification of the Asteraceae. More specifically, Chapter 3 outlines a published paper (Calabria et al., 2007) that covers the phylogenetic analysis of tribes of the Asteraceae based on phytochemical data. Chapter 4 covers the results from a meta-analysis of chemical data for the Asteraceae and represents a chapter in the conference volume for the International Compositae Alliance meeting in Barcelona, Spain in 2006 (Calabria et al., in press). This broad assessment of the chemistry of the Sunflower family provides a framework for

understanding the methodology and significance of studying saponin chemistry in the genus *Silphium*.

Chapter 5 describes the methods and results for the isolation and structural characterizaton of saponins from *Silphium*, focusing mainly on the nine new polyhydroxylated triterpene saponins isolated from *S. radula*, one of which showed significant inhibition of breast cancer cell proliferation (Calabria et al., 2008). Chapter 5.2 covers the saponin chemistry of the two prairie species, *S. terebinthinaceum* and *S. laciniatum*, including HPLC/ESI-MS profiles of both leaf and root extracts. Chapter 5.3 outlines structural and bioactivity data on saponins isolated from *S. integrifolium* and *S. morhii,* as well as an HPLC/ESI-MS analysis of intraspecific variation of saponins from different geographical populations of *S. perfoliatum*.

Chapter 6 describes the comprehensive LC-MS analysis used to examine the variation of triterpene saponins among all *Silphium* species. The medicinal importance and chemosystematic significance of these findings are also discussed. The final chapter, Chapter 7, summarizes my findings and suggests possible future studies of triterpene glycosides from *Silphium* and other members of the Asteraceae. The Appendices contain the majority of the NMR and MS data for the structures presented in Chapter 5 and 6, as well as a paper prepared on Dr. Mabry's Natural Product Chemistry career (Calabria and Mabry, 2007).

1.2 Why I Chose to Investigate Triterpene Saponins in *Silphium* for my Dissertation

This dissertation describes phytochemical and chemosystematic studies of one of the largest and most diverse plant families in the world, the Sunflowers. More specifically, most of my doctoral research has focused on the triterpene saponin chemistry of *Silphium* L. (Asteraceae), a small genus of resinous perennials native to North America and naturalized to parts of Europe where it is cultivated as a garden ornamental and herbal medicine. In 2000, the phylogenetic relationships of *Silphium* and subtribe Engelmanniinae were examined using DNA sequence data (both ETS and ITS). The results from these analyses indicated that the genus can be divided into two sections based on root type and growth form; section *Compositum* consists of 4 species and is characterized by taproots and scapiform growth, while section *Silphium*, with seven species, is distinguished by its fibrous roots and caulescent growth form. During the summers of 2002 and 2003, Ph.D. student Jeffery Williams and Dr. Jennifer Clevinger collected aerial parts of the eleven *Silphium* species from their native locations throughout the central and eastern United States. Thereafter, Dr. Mabry's phytochemistry group began a detailed investigation of the flavonoids and phenolic acids from *Silphium* species.

When I arrived at Dr. Mabry's lab in August of 2003 to begin my doctoral studies at UT-Austin, I became interested in this project because extracts of several *Silphium* species were widely used for herbal medicines by central and southeastern Native American tribes (Hamel and Chiltoskey, 1975), yet there had been few attempts to correlate phytochemical findings with their traditional herbal uses. One previous study

led to the isolation and characterization of eight major triterpene glycosides (saponins) from *Silphium* (Davidyants et al., 1984a), which showed blood cholesterol-lowering activity, as well as fungicidal properties (Syrov et al., 1992; Davidyants et al., 1997). These intriguing compounds are ubiquitous in the plant kingdom, although very little is known about their biological functions in plants or their distribution in the Asteraceae. This was surprising to me, given the large size of the Asteraceae (ca. 25,000 species) and its well-documented chemical diversity. Moreover, saponins are economically important compounds with a wide range of medicinal properties. For example, today, most commercial hormones and pain relievers are derived from saponins, including testosterone, progesterone and cortisone. Thus, in an effort to further characterize triterpene glycosides from *Silphium* and to determine some of their medicinal importance, I initiated an analysis of saponins in all species of *Silphium*.

Since the investigation of saponins was an entirely new area of research in the Mabry lab, the first step in my dissertation project was to gain instrumentation training and learn isolation and chemical techniques to characterize saponins. In August 2003, I began training with three post-doctorates (Dr. Malgorzata Wojcińska, Prof. Nabil El-Sayed, Prof. Ahmed Ahmed) in Mabry's lab on the extraction and isolation of saponins from *Silphium compositum*. In addition, I established connections with a number of leaders in the field of saponin chemistry, particularly, during my visit to the Institute of Soil Science and Plant Cultivation in Pulawy, Poland. While in Poland, I presented a poster at the International Conference on Plant Saponins (September 2004), and later worked in the laboratory of saponin specialist Prof. Wieslaw Oleszek to learn procedures

essential for the characterization of saponins. In the summer of 2005, one of Prof. Oleszek's very best post-doctorates, Irenuesz Kapusta, accepted a position in our lab to continue our investigations of saponins from *Silphium.* That summer we isolated nine new triterpene saponins from the leaves of the previously uninvestigated *S. radula*; all the aglycones were described for the first time. I mention that discovering new aglycones of this type is very rare, especially in a single species! Once the major saponins from *Silphium* were characterized, the cytotoxic activities of the isolated compounds were tested against human breast cancer cell lines.

As a result of my investigations of triterpene saponins from *Silphium,* I became interested in the chemosystematic value of these compounds as tools for understanding evolutionary and systematic aspects of the Asteraceae. I was surprised to find such a wide range of structural types and patterns of occurrences within a single genus and in a subtribe that has almost no chemical reports in the literature. While some species displayed a rich mixture of saponins, others contained only a few, structurally simple compounds. So I decided to overlay the chemical findings for *Silphium* with the most current DNA-based phylogenetic data to examine the patterns of occurrence of saponins in relation to the geographical and ecological distribution of the species. After reviewing the available literature on chemical data for the Asteraceae I found that Professor Vicente Emerenciano's group at the University of Sao Paulo, Brazil had created a large chemical database comprising ~400 skeletal types of terpenes, coumarins, flavonoids and polyacetylenes isolated from the Asteraceae. When I contacted Professor Emerenciano, he graciously granted me access to his database. Three months later we published the first

phylogenetic analysis of the entire Asteraceae based solely on chemical data (Calabria et al., 2007). Later that summer, I was invited to present a lecture and prepare a chapter on the chemistry of the Asteraceae for The International Compositae Alliance meeting in Barcelona, Spain. I asked Prof Emerenciano and his excellent graduate student, Marcus T. Scotti to co-author the chapter, which outlined the most current data available on the secondary chemistry of the Asteraceae in the context of the most up-to-date DNA-based phylogeny for the family (Calabria et al., in press).

Most recently, I have initiated a joint project with Dr. Paul Pare's research group at Texas Tech University involving the LC-MS analysis of species from *Silphium* and the closely related, *Lindheimera texana*. The main objectives of this project were to establish an LC-MS detection method for saponins in *Silphium* and to identify the major saponins in all *Silphium* species by comparison with the confirmed saponin standards isolated from *Silphium radula*. The data generated from this study were used to further characterize saponins from *Silphium* species and to target unknown compounds for future analysis. The results from the LC-MS analysis of extracts of *Silphium* species have implications for discovering new saponin structures from sources that were examined here for the first time. Moreover, the variation observed in saponin profiles may also yield insights into the evolution and diversification of chemical pathways for saponins in the Asteraceae and other plant families.

1.3 Research Objectives

The primary objective of my dissertation research was to develop and execute a standardized procedure for the extraction and identification of saponins from *Silphium* species in order to determine the precise chemical structures of the major saponins in these species. A variety of chromatographic techniques were employed for the isolation and separation of individual compounds, including solid-phase extraction (SPE), medium-pressure liquid chromatography (MPLC) and high-pressure liquid chromatography (HPLC) with UV and MS detection. Structures of saponins isolated from *Silphium* were determined primarily by NMR, ESI-MS and tandem MS/MS analysis, working with the NMR and MS facilities at UT-Austin.

A second major objective of my research was to establish a comprehensive HPLC/ESI-MS method for the metabolic profiling of saponins in *Silphium*. This type of method has allowed for the tentative identification of a large number of triterpene saponins from all species examined that would not have been possible using the relatively time consuming isolation and structure determination procedures typically employed for saponins. The data generated from this analysis were used to examine the qualitative variation and distribution of saponins across the entire genus. Moreover, the chemosystematic value of saponins was evaluated as a starting point for future studies on the triterpene chemistry of the Asteraceae.

Chapter 2: Background

A. *Silphium* L. (Asteraceae)

This section of Chapter 2 provides an introduction to the main subject of this dissertation, *Silphium*. A historical overview of the taxonomic and systematic treatment of *Silphium* is presented and in addition, the natural history, ethnobotany and chemistry of the genus are outlined.

2.1 Natural History

Silphium L. (Asteraceae) is a small genus of perennial sunflowers found throughout the prairies and woodlands of central North America (Cronquist, 1980). It is difficult to provide a general description of the growth form of *Silphium* because of the great variation within species, depending on climatic conditions and between species growing in different geographical and ecological areas. However, there are some morphological and anatomical characteristics that are common to all *Silphium* species, such as their coarse, resinous leaves and stems and medium to large-sized, yellow flowering heads (except for *S. albiflorum*, which has white flowers) where only the ray flowers are pistillate and the disk flowers are functionally staminate (Figure 2.1). *Silphium* can be recognized as a member of the Heliantheae from its carbonized cypselae and paleate receptacles. *Silphium* belongs to the subtribe Engelmanniinae due to its

Figure 2.1. Resinous gum exudes from the stem of *S. laciniatum*. (Image courtesy of Kenneth Robertson)

staminate disc flowers, radially flattened ray cypsela and presence of alternate leaves in some taxa, but lacks the cypsela complex present in other members of this subtribe. Instead, the cypsela in *Silphium* is fused only to the subtending phyllary; the other members of Engelmanniinae have cypsela fused also to 2-4 disc flowers.

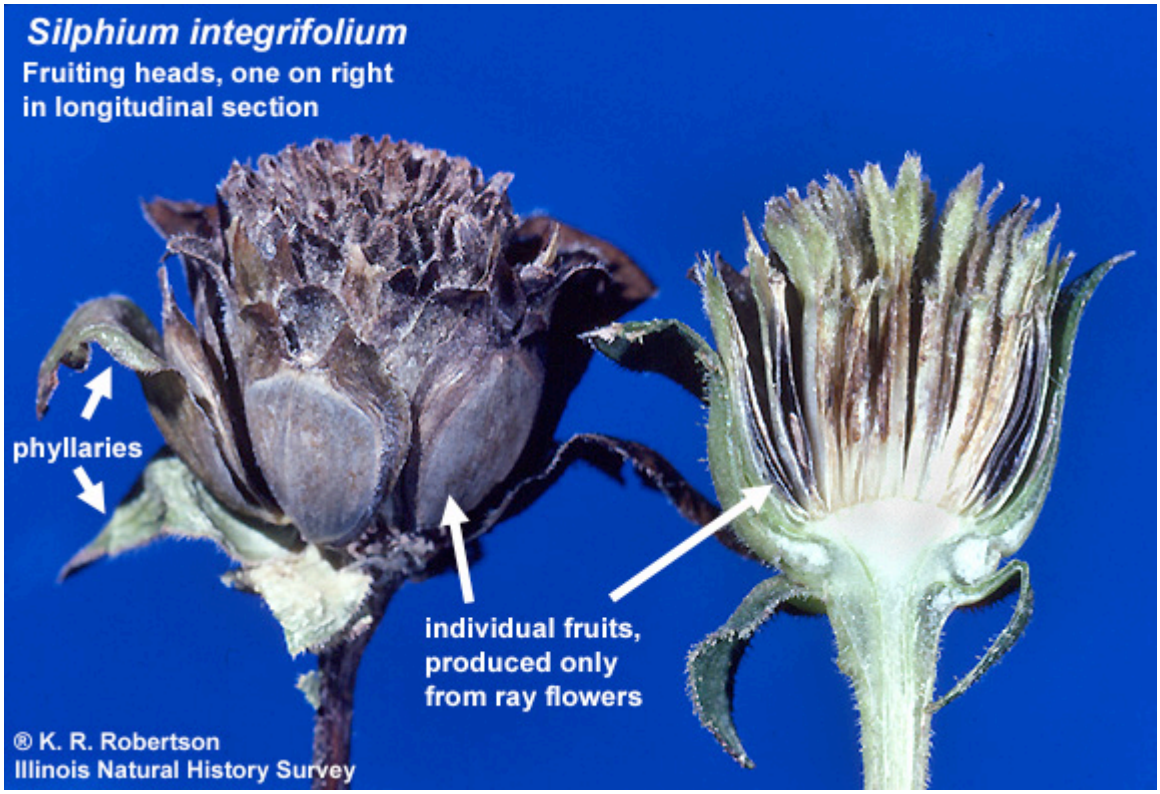


Figure 2.2. Longitudinal section of fruiting head of *S. integrifolium* illustrating the mature fruits and phyllaries subtending the individual fruits. (Image courtesy of Dr. Kenneth Robertson)

Like all other sunflowers, the capitulescence (a special term used in Asteraceae to describe a group of associated heads, analogous to an inflorescence) of *Silphium* are cymose and the capitula are heterogamous and campanulate to hemispherical in outline. The receptacle is flat to slightly convex and the shape of the pales (chaffy bracts), disc flowers and ray flowers are unvarying throughout the genus. The bracts subtending and peripheral to the ray flowers, called phyllaries, are arranged in two or three graduated

series, where the outermost are broad and foliaceous, while the innermost are narrower and membranaceous (Figure 2.2).

The ray flowers are pistillate and fertile, the ovaries imbricate in 2-3 series, herbaceous or partially memberanous-charttaceous; the ray cypsela are glabrous, strongly flattened parallel to the involucral bracts with winged margins. There is typically no pappus or a pappus containing 2 awns confluent with the wings of the cypsela. The disc flowers are perfect but functionally staminate, with unbranched stigmas and a linear and elongate sterile disc ovary.

The leaves of *Silphium* can be variously arranged as opposite, alternate or whorled and sometimes an individual or a population will exhibit more than one arrangement at different heights on the stem, leading taxonomists to wrongfully delineate a new species, when, in fact, they were observing population-level morphological variation. All *Silphium* species have teeth at least on their lowermost leaves; in *Silphium* section *Compositum* the leaves can be toothed in a pinnatfid or palmatifid fashion or lobed to deeply lobed. The variation in leaf morphology has lead to confusion regarding hybridization between species in *Silphium* Section *Compositum*, such as *S. terebinthinaceum* and *S. laciniatum* (see Chapter 5.2 for more on this topic). Most species in *Silphium* section *Silphium* have lanceolate to ovate leaves, with several distinct shapes in certain species. For example, the well-known *S. perfoliatum* has a distinct connate-perfoliate leaf petiole that creates a cup, which catches rainwater, giving its common names, Indian Cup or Cup plant.

Silphium has a primary chromosome number of *n*=7 while *Lindheimera*, the sister genus to *Silphium*, has a base chromosome number of *n*=8. Furthermore, an individual of *Silphium* will occasionally have one extra unpaired or two paired chromosomes (Smith, et al., 1992; Keil et al., 1988; Parfitt, 1981). These data together suggest that the chromosome number in *Silphium* is the result of an anueploid reduction event.

Several researchers have reported on the karyotyping of chromosomes from *Silphium* species (Taylor 1926; Fisher and Cruden, 1962; Settle, 1967). There were some interesting results from these studies on reciprocal translocation between chromosome pairs in *Silphium*, where translocation events between chromosome 1 and 7 could be categorized into three groups: A, B and C. Species in Group C all belong to the section *Compositum*, whereas all of the species from group A and B belong to section *Silphium*. The results from this analysis of chromosome translocations are congruent with the results of Dr. Clevinger's revision of *Silphium* (see discussion in the next section of this chapter).

Silphium occurs in a variety of habitats including prairies, open forest and meadows but most frequently in areas created by man-made disturbances, such as roadside ditches, cemeteries, highway right-of-ways and along railroad tracks. Thus, despite much of *Silphium*'s original prairie habitat being degraded as a result of expanding human populations, this resilient sunflower has learned to exploit new habitats created by humans. *Silphium* can be found as far north as Canada and as far south as central Florida and central Texas. Any further south would not provide a freeze during the winter months, which is crucial for *Silphium* seeds to germinate.

Not much is known about *Silphium*'s seed dispersal methods or pollination. Because the cypselae of *Silphium* are large and lack both animal and wind-dispersal structures, such as the feather-like fiber parachutes of dandelions (*Taraxacum officinalis)*, seedlings nearest the flowering stems of a neighboring plant are most likely its progeny. One study showed estimates of the average distance for dispersal in *S. laciniatum* was just 1.0m from the parent plant (Pleasants and Jurik, 1992). Garden observations of *Silphium*'s dispersal range reinforce the concept that *Silphium* seedlings are often found growing in close proximity to the parent plant (Clevinger, 1999). Although there are no explicit studies on the pollination of *Silphium*, incidental observations suggest honey bees (*Apis mellifera*) are the major pollinators; for example, the maximum recorded foraging distance for bees is 13.7 km, but the average distance can be quite small (a few hundred meters) (Gary, 1992; Winston, 1987). The limited ranges for both seed dispersal and pollinators have created a situation in which populations are very localized and spatially isolated from one another; this may have contributed to the findings that several species of insects are specialist feeders of *Silphium* species. For example, *Okanagana balli* (Prairie cicada) grubs feed on the large taproot of *Silphium* species, while *Rynchites* sp. (*Silphium* beetle) and its larvae feed on the flower heads and stems. The oligolectic aphid *Iowana frisoni* sucks the juices from the flowering stems (Hilty, 2007). The larvae of *Antistrophus rufus* and *Antistrophus minor* (gall wasp) feed within the stems, forming galls that are not visible from the outside (see section on *S. terebinthinaceum* for more details). Moreover, they attract the hyperparasitic wasp *Eurytoma lutea*, whose larvae feed on these gall formers. Similarly, the larvae of *Mordellistena aethiops* (Tumbling

Flower Beetle sp.) feed within the stems, while the adults may eat the flowers Tooker and Hanks, 2004).

2.2 Taxonomic History and Current Phylogenetic Placement

Linnaeus was the first taxonomist to recognize *Silphium* as a genus in his work *Species Plantarum* (1753). He described five species, two of which were later transferred to the genus *Heliopsis*. Since his original classification, numerous species have been added to *Silphium*. Small (1933) described 33 species based on leaf, stem, capitulum and phyllary morphology. Perry (1937) included just 23 species in her treatment of *Silphium*. The next major revision to *Silphium* was conducted by Fisher and his team of graduate students at Ohio State University, where they separated the genus into eight morphologically distinct groups containing a total of only 16 species, based on a series of chromosomal, morphological and hybridization studies; however, no formal revision of the genus was ever published by Fisher and co-workers. Then in 1980, Cronquist published a revision of *Silphium,* including 15 species, as a part of the Asteraceae volume of the Flora of the Southeastern USA. One additional species from Kentucky and Tennessee, *S. wasiotense*, was added to *Silphium* by Medley (1989), who hypothesized that *S. wasiotense* was most closely allied to *S. brachiatum* and *S. morhii*.

More recently, Clevinger (1999) and Clevinger and Panero (2000) published a revision of *Silphium* based on molecular, morphological and chromosomal data, including a detailed inspection of herbarium specimens. The phylogenetic relationships of *Silphium* and subtribe Engelmanniinae were examined using DNA sequence data (both

ETS and ITS). The results from these analyses indicated that the genus can be divided into two sections based on root type and growth form; section *Compositum* consists of 4 species and is characterized by taproots and scapiform growth, while Section *Silphium*, with seven species, is distinguished by its fibrous roots and caulescent growth form (Figure 2.3). Further combinations by Clevinger (2004) evaluated the status of several species and her results indicated that the species studied were actually varieties of either *S. asteriscus* or *S. radula*. The most current circumscription of *Silphium* includes one additional species from Alabama, *Silphium glutinosum* Allison, bringing the total number of species in the *Silphium* complex to twelve (Clevinger, 2006).

The findings discussed in this dissertation are based on the eleven species classification system proposed by Clevinger in 1999 (Figure 2.3). Because the large-scale isolation of triterpene saponins from *Silphium* species was completed in 2006 the additional species, *S. glutinosum* (Clevinger, 2006), was not investigated in this dissertation research.

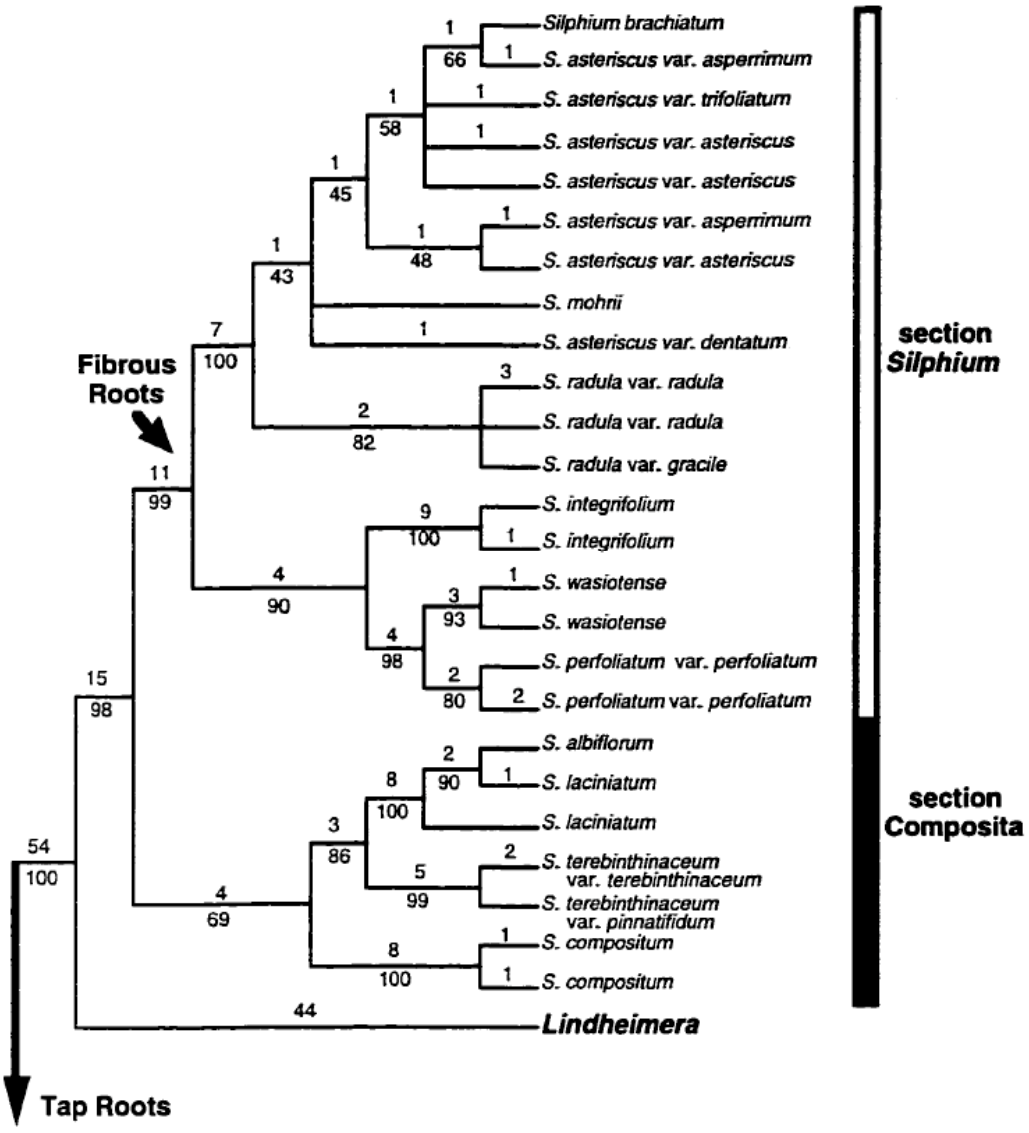


Figure 2.3 .The phylogenetic tree for *Silphium* and the related genera *Lindheimera* based on ITS and ETS sequence data. This figure represents the eleven species classification system used in this dissertation to examine the triterpene saponin chemistry of the *Silphium*, before *S. glutinosum* was recognized as a species in 2006. (Figure courtesy of Dr. Jennifer Clevinger, Clevinger, 1999)

2.3 Ethnobotany

There are several reports on the medicinal uses of *Silphium* by various Native American tribes, including the Cherokee, Chippewa, Dakota, Iroquois, Ojibwa , Omaha, Meskwaki, Pawnee, Ponca and Winnebago (Shemluck, 1982; Hamel and Chiltoskey,

1975; Moerman, 1991; Moerman, 2003). In addition, early American herbalists reported on *Silphium* for the treatment of a variety of ailments in several *Materia Medica* texts (Cook, 1869; Felter and Lloyd, 1898; Grieves, 1971). Table 2.1 provides a summary of information on the medicinal uses of *Silphium* species. Additional information on the Native American tribes and the plant parts used are also provided.

Medicinal uses	**Species**	**Plant part**	**Native American Tribe(s)**
Abortifacient	*S. compositum*	N/A	Chippewa
Alterative	*S. perfoliatum*	root infusion	Cherokee
Analgesic	*S. perfoliatum* *S. integrifolium*	root infusion	Cherokee, Meskwaki, Chippewa, Ojibwa, Ponca, Winnebago
Anti-diarrheal	*S. laciniatum*	root infusion	Ojibwa
Anti-emetic	*S. perfoliatum*	root decoction	Meskwaki, Ojibwa
Anti-hemorrhagic	*S. perfoliatum*	root infusion	Chippewa
Anti-rheumatic	*S. perfoliatum*	root infusion	Ojibwa, Omaha, Ponca, Winnebago
Anti-spasmodic	*S. laciniatum*	root infusion	Chippewa
Asthma	*S. laciniatum*	root infusion	Chippewa
Chewing gum	*S. laciniatum*	stem resin	Dakota, Winnebago, Ponca, Pawnee, Omaha
Diuretic	*S. integrifolium*	leaves, roots	Meskwaki
Diaphoretic	*S. perfoliatum*	roots	Meskwaki
Emetic	*S. laciniatum*	root decoction	Meskwaki, Iroquois
Emmenagogue	*S. perfoliatum*	root decoction	Chippewa
Enlarged spleen	*S. perfoliatum*	root smoke	Chippewa
Febrifuge	*S. perfoliatum*	root infusion	Cherokee
Head cold	*S. perfoliatum*	root smoke	Omaha, Ponca
Hepatic	*S. perfoliatum*	root smoke	Cherokee
Internal bruising	*S. perfoliatum*	root infusion	Cherokee
Nervine	*S. laciniatum*	root infusion	Meskwaki
Neuralgia	*S. perfoliatum*	root smoke	Omaha
Stimulant	*S. compositum* *S. laciniatum*	stem resin	Cherokee
Tonic	*S. laciniatum*	root infusion	Pawnee
Vermifuge (horses)	*S. laciniatum*	root infusion	Dakota, Omaha

Table 2.1. Summary of the ethnobotanical uses of *Silphium* species by Native American tribes.

Based on the ethnobotanical uses of *Silphium* by early Americans, Kindscher and co-workers (1998) collected *S. perfoliatum* and *S. laciniatum* and conducted biological screening on aqueous and methanolic extracts to identify new anti-HIV and anti-cancer leads. The results from the screening indicated organic extracts of the leaves, stems and roots of *S. perfoliatum* and *S. laciniatum* exhibited moderate activity in an anti-cancer screening (LC_{50} > or = 20% of the cell lines) and aqueous extracts of *S. laciniatum* showed strong anti-HIV activity (LC_{50} > 50% compared with the control), while the *S. perfoliatum* organic extract showed moderate activity (LC_{50} <50% of control) in the anti-HIV assay (Kindscher et al., 1998).

2.4 Chemistry

Silphium, like most other genera in the Asteraceae, is characterized by its diverse chemistry, including flavonoids (El-Sayed et al., 2002; Williams, 2006), phenolic acids (Kowalski and Wolski, 2003), essential oils (Kowalski et al., 2005), sesquiterpenes (Bohlmann and Jakupovic, 1979; 1980), diterpenes (Pcolinski et al., 1994). In addition, *Silphium* species also contain an abundance of oleanolic-type triterpene saponins (Davidyants et al., 1984a,b,c; Davidyants et al., 1985; Davidyants et al., 1986; Calabria, et al., 2008). Although this dissertation focuses on the triterpene saponins of *Silphium*, the literature data on all other groups of secondary metabolites is presented in this section. Saponins are discussed in greater detail in Section B (page 23).

2.4.1 Diterpenes

Pcolinski (1994) isolated a series of labdane-type diterpenes, namely chlorosilphanol-A, B and C and silphanepoxol, from *S. perfoliatum* whose structures were determined by NMR, MS and X-ray crystallography. Chlorine-containing diterpenes are not common in the plant kingdom and this report was the first to assign the absolute stereochemistry of this group of labdane diterpenes. Other structurally interesting diterpenes have been isolated from the roots of *S. perfoliatum*, including isocomene, modhephene and β-isocomene (Paquette and Leone-Bay, 1983).

2.4.2 Sesquiterpenes

A few reports on sesquiterpenes from *Silphium* have been published. Bohlmann and Jakupovic (1980) reported three new sesquiterpenes. A highly condensed tricylcopentanoid sesquiterpene, known as silphinene, was characterized by Paquette and Leone-Bay (1983). While analyzing the essential oil composition of *S. perfoliatum*, Kowalski and Wolski (2005) found the dominant constituent of the leaves to be a bicyclic sesquiterpene, caryophyllene oxide and the monocyclic sesquiterpene germacrene D. In the rhizome oil, the most abundant components were the tricyclic sesquiterpenes, 7-β-H-silphiperfol-5-ene, isocomene, modhephene and 7-α-H-silphiperfol-5-ene (Kowalski and Wolski, 2005).

2.4.3 Flavonoids

Our research group conducted one of the first investigations of flavonoids from *Silphium perfoliatum,* which resulted in the discovery of two apiose-containing kaempferol triosides and nine known flavonoids (El-Sayed et al., 2002). Two additional, novel apiose-containing and six known flavonol glycosides were later isolated from *S. albiflorum* and were further characterized through LC-MS with post-column manganese complexation (Wojcinska et al., 2006).

A thorough investigation of the flavonoid chemistry of all *Silphium* species was recently carried out by Williams (2006), confirming the diversity of flavonoids in this genus and reporting a few additional novel flavonoid triglycosides. The results from these analyses indicated that most flavonoids found in *Silphium* are mono- or disaccharide derivatives of the flavonols quercetin, isorhamnetin and kaempferol. Quercetin and isorhamnetin appeared to be the most common flavonols present. The cytotoxic activities of the new triglycosides of quercetin, kaempferol and isorhamnetin were tested against human breast cancer cell line MCF7. Results showed that three of the new flavonoid triglycosides [quercetin 3-*O*-α-L-rhamnosyl (1-6)-*O*-β-D-galactopyranoside 7-*O*-β-L-apiofuranoside, quercetin 3-*O*-β-L-apiofuranoside 7-*O*-α-L-galactosyl (1-6)-*O*-β-D-rhamnopyranoside and kaempferol 3-*O*-β-D-apiofuranoside 7-*O*-α-L-rhamnosyl (1-6)-*O*-β-D- (2-*O*-*E*-caffeoylgalactopyranoside)] decreased cell proliferation in a statistically significant manner in comparison with the control (Williams, 2006).

Recently, Valant-Vetschera and Wollenweber (2007) examined exudate flavonoid diversity in members of the Asteraceae, including *Silphium*. Their analysis of

S. laciniatum resulted in the detection of several flavonoid aglycones including derivatives of kaempferol, quercetin, quercetagetin, eriodictyol and two flavonoid glycosides: kaemperfol 3-glucoside, quercetin 3-glucoside. *S. terebinthinaceum* yielded kaempferol 3-glucoside, quercetin 3-rhamnoside, quercetin 3-glucoside and quercetin 3-rhamnoglucoside. Interestingly, their results showed that this species lacked aglycones in the exudates, which is quite uncommon in the Asteraceae (Valant-Vetschera and Wollenweber, 2007).

2.4.4 Phenolic acids

Wojcinska and Drost-Karbowska (1998) conducted a detailed study on the phenolic acids of *Silphium perfoliatum* flowers. A further LC-MS analysis of the phenolic acid chemistry of *Silphium* showed the presence of large amounts of hydroxycinnamic acids; specifically, p-coumaric, hydrocaffeic and isoferulic acids were the most abundant phenolics found in *Silphium* (Williams, 2006).

2.4.5 Monoterpenes

Monoterpenes are volatile terpenes with low molecular weights and are typically the main components of the essential oils in plants. For *Silphium*, there are several literature reports on the ecological implications of volatiles from *S. terebinthinaceum* and *S. laciniatum* in relation to the tritrophic interactions with the gall wasp, *Antistrophus* and the parasitoid *Eurytoma.* The larvae of the gall-forming wasp, *Antistrophus* inhabit the pith and cambium of stems of the tap-rooted prairie perennials *Silphium terebinthinaceum and Silphium laciniatum* (Tooker et al., 2002; Tooker and Hanks 2004;

Tooker et al., 2005, Tooker and Hanks, 2006). The gall wasp *A. rufus* alters the ratios of host plant monoterpenes, namely a blend of α and β-pinene and camphene, providing an olfactory signal that males use to detect their mates. The ratios of these monoterpenes in the blend are critical for host recognition as demonstrated by several field, common garden and laboratory experiments (Tooker et al., 2005). This same blend of volatiles emitted by the plant also attracts the natural enemies of *Antistrophus,* the parasitoid *Eurytoma*, thus increasing the fitness of the host plant by killing the gall wasp. This phenomenon is discussed in greater detail in Chapter 5.2.

The chemical composition of the essential oils in leaves, inflorescences and rhizomes of *Silphium perfoliatum* were examined by Kowalski and Wolski (2004). Over 60 mono-, di and sesquiterpenes make up the composition of the essential oils *of S. perfoliatum* and are too numerous to list in full here; however, the dominant monoterpene in the essential oil mixture is α-pinene, which concurs with Tooker et al. (2005) studies on volatiles present in *S. laciniatum.*

The content of the essential oil from leaves and inflorescences of *Silphium integrifolium* was also examined by Kowalski and co-workers (2005) and they found the main monoterpene constituents in the leaves and inflorescence were α-pinene, trans-verbenol, bornyl acetate and allo-aromadendr-9-ene.

2.4.6 Triterpenes

The first report on triterpenic compounds from *Silphium* described the isolation and structural characterization of a series of oleanolic acid saponins (silphioside A, B, C and E and glycosides F and G) from the aerial parts of *S. perfoliatum* (Davidyants et al., 1984 a, b, c). Silphioside B is one of the most abundant saponins found in *Silphium perfoliatum* (Figure 2.4).

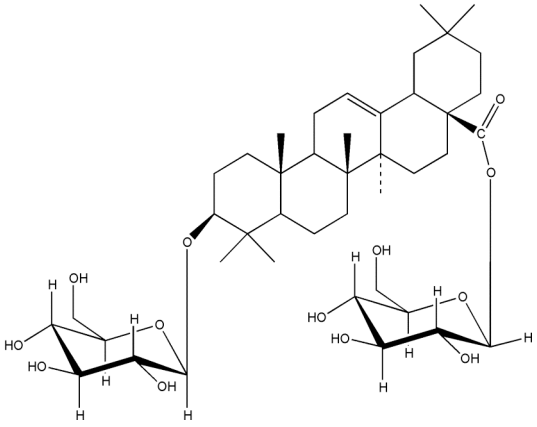


Figure 2.4. An example a saponin typically found in *Silphium perfoliatum*, silphioside B, first isolated by Davidyants et al., 1984a.

Several subsequent reports were published describing a range of biological activities for these compounds; for example, a mixture of saponins isolated from *S. perfoliatum* (silphiosides A, B, C D and glycosides F and G) showed blood cholesterol-lowering activity when administered orally in both normal and hyperlipidemic rats (Syrov et al., 1992). Another study described the anti-fungal properties of triterpene saponins from *S. perfoliatum* where a mixture of the saponins significantly inhibited the growth of *Dhreslera graminea* and inhibited both mycelial growth and spore formation of *Rhizopus nodosus* and *R. nigricens* (Davidyants et al., 1997). More recently, the growth-regulating effects of saponins from *S. perfoliatum* were reported where a mixture of silphiosides showed cytokinin and auxin-like activities on seed germination, hypocotyl elongation and root formation (Davidyants, 2006).

In addition to *Silphium perfoliatum*, there has been an increasing interest in European-cultivated species of *Silphium* as commercial sources for oleanolic and ursolic-type triterpenes. For example, Kowalski (2007) published a qualitative and quantitative evaluation of triterpene aglycones from saponin fractions of three species of *Silphium*, namely *S. perfoliatum, S. trifoliatum* and *S. integrifolium*, and compared the findings for the *Silphium* species with those for *Panax quinquefolium* roots and *Calendula officinalis* flowers using HPLC-PDA/ESI/MS. The results from this study revealed the presence of 36 oleanolic and ursolic acid saponins in leaves, stems and roots of the three *Silphium* species examined. However, Kowalski's study did not include authentic standards isolated from *Silphium* species and only one of the 36 compounds was fully identified.

Our research group has recently reported on the isolation and structural elucidation of nine new polyhydroxylated pentacyclic oleanene and ursene-type triterpene saponins from methanolic extracts of leaves and stems of *Silphium radula* Nutt. (Calabria et al., 2008) (see Figure 5.1, Chapter 5.1 for structures). The cytotoxic activities of the isolated compounds were tested against human breast cancer cell line 25 MDA-MB-231, and the results showed that saponin **I** decreased breast cancer cell proliferation in a statistically significant manner at 25 μg/ml.

B. Saponins

This section of Chapter 2 provides an overview of saponins, a widespread and economically important group of natural products. Here, the important features of saponins are outlined, as well as their historical and medicinal significance, including

information on their basic structures, biosynthesis, functions and bioactivities as they relate to *Silphium*, which has been established as a rich source of triterpene glycoside-type saponins (Davidyants et al. 1984a, b, c; Kowalski, 2007; Calabria et al., 2008).

2.5. Historical Overview of Saponins

Plants containing saponins have been recognized for centuries not only for their detergent-like properties, reflected from the Latin-derived *sapo* (soap), but also for their value in folk medicines. Much of the early research, from the 1930's through the 1960's, on saponin-containing plants was motivated by the search to discover steroidal saponins that could serve as precursors of hormones and other steroid drugs. Prior to the 1940's, clinical studies on steroid drugs were severely hindered by the lack of sufficient quantities of compounds for testing. Then in 1943, Russell Marker developed a method for synthesizing progesterone from diosgenin, the aglycone of a saponin naturally occurring in Mexican wild yams *(Dioscorea mexicana)*. Today, most commercial hormones and pain relievers, including testosterone, progesterone and cortisone, are derived from plant steroidal saponins. Despite the commercial interest and important biological activities of saponins, much remains unknown about their distribution, medicinal properties and functional importance in plants; even less is known about environmental factors that influence saponin production in plants.

2.6 Structures and Biosynthesis of Saponins

Structurally, saponins consist of a triterpenoid or steroid aglycone with a sugar chain attached to the aglycone most commonly at the C-3 hydroxyl position and may

have additional glycosidic groups at other positions. The glycosidic groups often consist of more than one sugar and can be modified by acetyl, methoxyl or other functional groups. Figure 2.5 provides an example of steroidal and triterpene-type saponins.

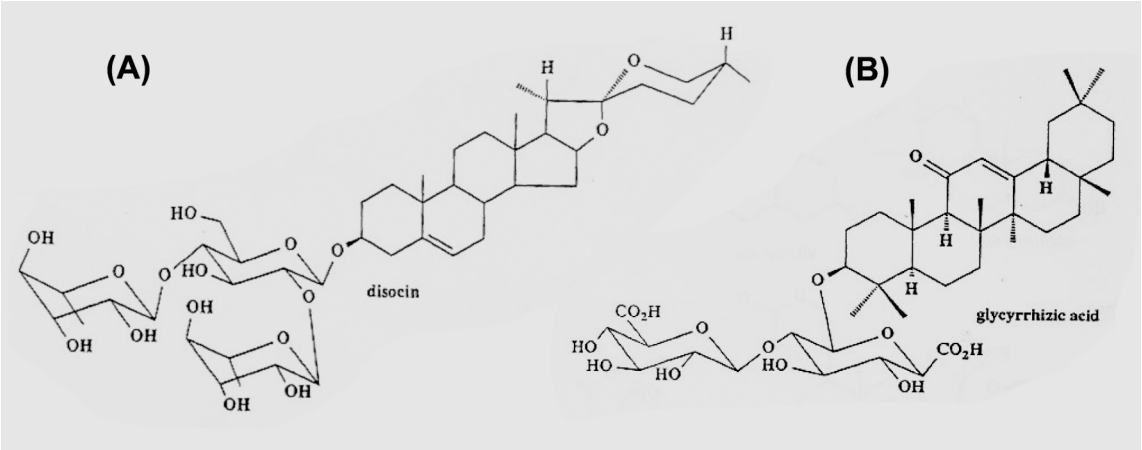


Figure 2.5. Examples of steroidal (A) and triterpene (B) saponins. Dioscin (A) (aglycone diosgenin), extracted from wild yam (*Dioscorea villosa*), is used as starting material in the semisynthesis and microbial transformation of steroidal hormones, such as progesterone and corticosteroids. Glycyrrhizic acid (B) is the active principle of liquorice root, and a powerful sweetener (30-50 times as potent as sucrose). These images were provided courtesy of Prof. David Seigler.

Triterpenoid saponins are synthesized from mevalonic acid via the isoprenoid pathway as a normal part of growth and development in many plant species. Three C-5 isoprene units are joined together in a head-to-tail manner resulting in a 15 carbon molecule, farnesyl pyrophosphate. Two farnesyl pyrophosphate units are then linked tail-to-tail to give squalene (C-30) (Figure 2.6). Next, squalene is reduced to 2, 3-oxidosqualene which can be cyclized to any number of different triterpene skeletons. This cyclization reaction is considered the first committed step in triterpenoid biosynthesis and marks the branching point with the sterol biosynthetic pathway; sterols are very important membrane constituents and serve as precursors for hormone biosynthesis.

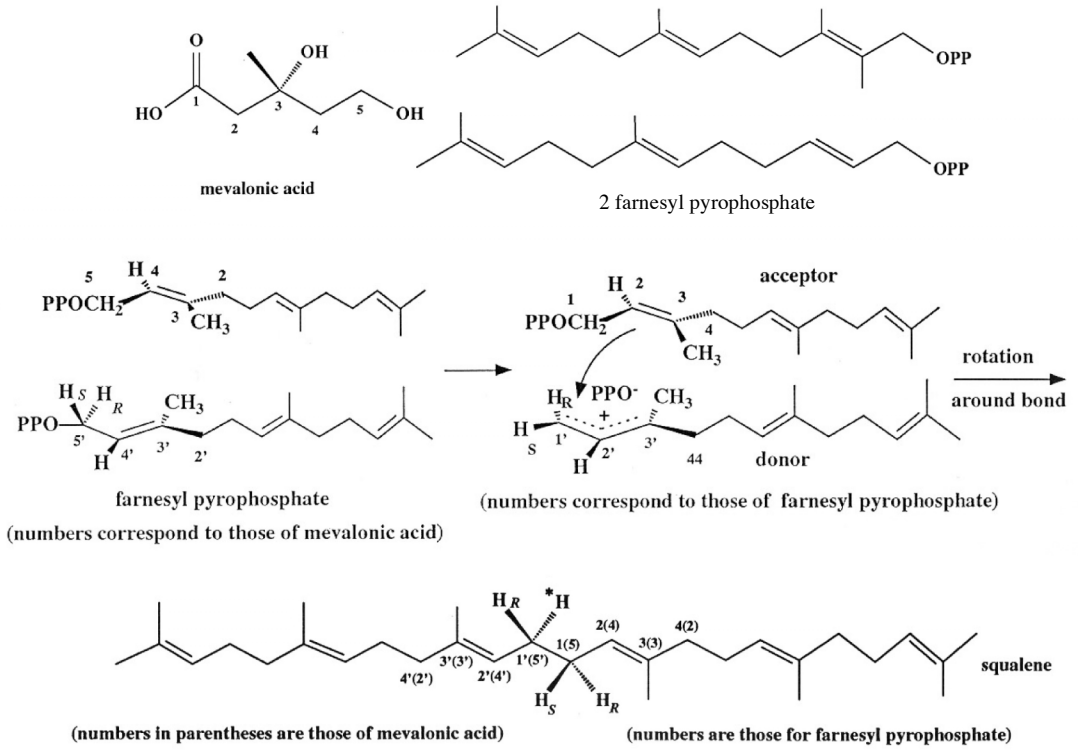


Figure 2.6. Mevalonic acid (top left) is the immediate precursor of the C5 isoprene unit, which is the basic building block for all terpenes. The C30 compound known as squalene (bottom) is derived from two farnesyl pyrophosphate units (top right), which undergo a series of condensation reactions (center reaction sequence) to produce the unusual head-to-head arrangement of squalene. This image was provided courtesy of Prof. David Seigler.

Figure 2.7 provides an illustration of the biosynthetic pathway leading to pentacyclic triterpenes. The triterpenoid backbone then undergoes modifications (oxidations, substitutions, glycosylation), mediated by cytochrome P450-dependant monooxygenases, glycotransferases and other enzymes (Achine et al., 2005). The specific products and attachments to the triterpene skeleton can vary widely between plant species. To date, several triterpene synthases from plants have been structurally characterized (Iturbe-Ormaetxe et al., 2003; Xiang et al., 2006; Basyuni, et al., 2006). However, because of the lack of pathway intermediates for biochemical studies and due

to the complexity of the saponin structures, there is still much to learn regarding specific enzymes and biochemical pathways involved in saponin biosynthesis.

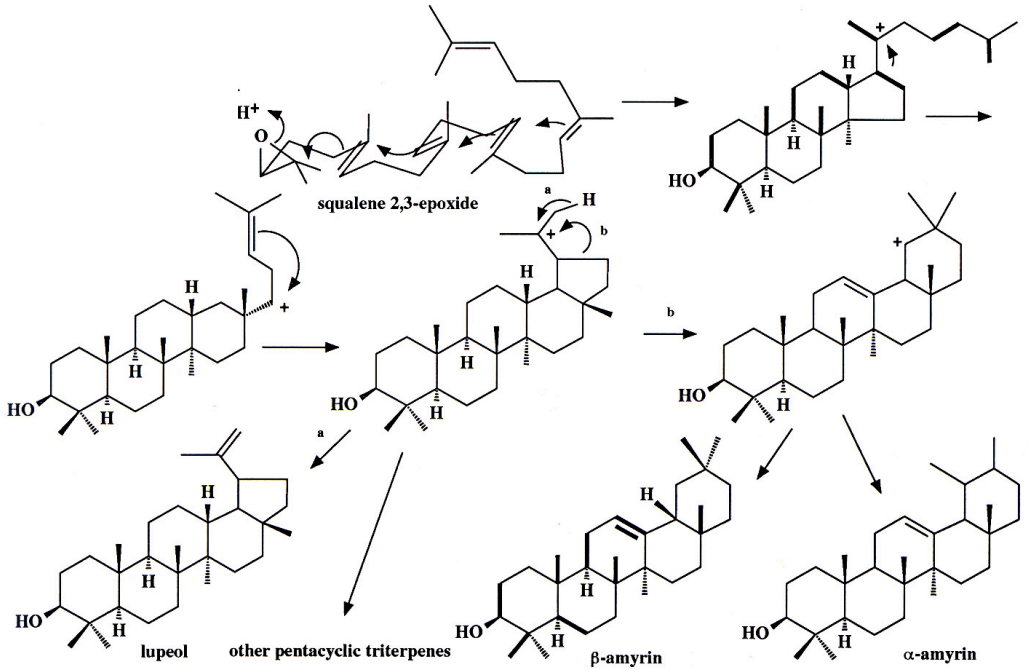


Figure 2.7. A diagram representing the biosynthetic pathway leading to the formation of pentacyclic triterpenes. Squalene is oxidized to produce squalene 2,3 epoxide. The triterpenoid backbone then undergoes modifications (oxidations, substitutions, glycosylation), leading to a diversity of compounds including membrane sterols, brassinosteroids, phytohormones, cuticular waxes and saponins. This image was provided courtesy of Prof. David Seigler.

2.7 Functions of Saponins in Plants

Many triterpene saponins possess anti-bacterial and anti-fungal properties and have well-established roles in protecting plants against pathogenic attacks (Papadopoulou et al., 1996). In addition, some of these molecules have other suggested functions in plants. For example, saponins from pea (*Pisum sativum*) have been proposed to regulate gravitropism. In 2001, Rhaman and co-workers published a study in *Plant Physiology*

showing that chromosaponin, a pyronyl-triterpenoid isolated from pea, specifically interacts with the AUX1 protein in regulating the intracellular levels of auxin in *Arabidopsis* roots. The AUX1 gene encodes an auxin influx carrier, an amino acid permease-like protein embedded in the plasma membrane of cells involved in auxin transport. The results from this study showed that chromosaponin I (CSI) inhibits gravitropism, as well as auxin influx and ethylene-mediated growth responses in wild-type roots and restores these actions in *aux1-7* mutants.

To understand the mechanism by which CSI blocks ethylene responses, Rhaman et al. (2001) used CTR1 mutants; these mutants constitutively display the ethylene triple response in the absence of ethylene. The results from this experiment showed that CSI inhibits ethylene signaling at or downstream of CTR1. The authors propose that CSI reduces the intracellular level of auxin in roots by inhibiting auxin uptake, resulting in resistance to ethylene. In *aux1-7* mutants roots CSI increases intracellular levels of auxin levels by stimulating auxin uptake, resulting in restoration of the ethylene response (Rhaman et al., 2001).

Ohana and co-workers (1998a) reported on another triterpenoid saponin (GTS) isolated from peas. GTS is a potent inhibitor of diguanylate cyclase (dgc), a key regulatory enzyme of cellulose biosynthesis in the bacterium *Acetobacter xylinum*. In addition, a very similar saponin (with identical biological activity) was also isolated from *A. xylinum*. The evidence from this study strongly suggests that certain aspects of cellulose synthesis in bacteria and higher plants may be regulated by saponins (Ohana et al., 1998b).

2.8 Medicinal Actions and Bioactivities of Saponins

Saponins are bioactive compounds with a wide range of medicinal properties including hypocholesterolemic, anticarcinogenic, anti-inflammatory and anti-oxidant activity in animals and *in vitro* models (Rao and Gurfinkel, 2000). Saponins are found in a number of plant-derived foods such as legumes, and are present in many medicinal plants, including ginseng. Here, the most important activities as they relate to the study of triterpene saponins from *Silphium* and other members of the Asteraceae are highlighted.

2.8.1 Anti-cancer Activity

There have been numerous reports of triterpenes and their glycosides that exhibit anti-proliferative and anti-tumor activities from members of the Asteraceae. For example, astersedifolioside A-C from *Aster sedifolius* were tested for their anti-proliferative effects against a transformed thyroid cell lines and were found to arrest the cell growth in a dose dependent manner (Corea et al., 2004). In addition, evaluation of the anti-tumor properties of 15 triterpene alcohols isolated from *Chrysanthemum* flowers (Asteraceae), revealed that all compounds clearly suppressed the promoting effects of TPA on skin tumor formation in mice (Ukiya et al., 2002). In addition, six of these compounds also exhibited cytotoxic effects on human cancer cell lines.

2.8.2 Anti-inflammatory

Triterpene alcohols from the flowers of a variety of Asteraceae genera, including *Calendula, Cosmos, Chrysanthemum, Helianthus* and *Matricaria* showed marked

inhibitory activity against TPA-induced inflammation (Akihisa et al., 1996; Ukiya, 2001). Furthermore, a study testing the anti-histaminic and anti-eicosanoid effects of oleanolic and ursolic acid type triterpenes from *Helichrysum picarardii* (a medical herb from the Asteraceae) showed that these compounds were responsible for the anti-inflammatory properties of this medicinal plant (Santos Rosa et al., 2007)

2.8.3 Cardiovascular Activity

Although there have been a variety of studies conducted on the effects of saponins on the cardiovascular system, the most relevant to *Silphium* is hypocholesterolemic activity, where saponins act by decreasing the intestinal adsorption of cholesterol resulting in the lowering of plasma and hepatic cholesterol levels. It is possible that the mechanism of this action is related either to the saponins ability to bind cholesterol in the gut, thus inhibiting the direct absorption of cholesterol, or to the saponins' effect on cholesterol metabolism by interfering with the circulation of bile acids. A mixture of saponins isolated from *S. perfoliatum* (silphiosides A, B, C D and glycosides F and G) showed blood cholesterol-lowering activity when administered orally in both normal and hyperlipidemic rats (Syrov et al., 1992).

2.8.4 Anti-viral, Anti-microbial and Anti-fungal Activities

Many saponins display activities related to their role in plant defense against pathogens. As mentioned before, Davidyants et al. (1997) evaluated the anti-fungal properties of triterpene saponins from *S. perfoliatum*, where a mixture of the saponins

significantly inhibited the growth of *Dhreslera graminea* and inhibited both mycelial growth and spore formation of *Rhizopus nodosus* and *R. nigricens*. (Davidyants et al., 1997). In addition, saponins isolated from *Calendula officinalis*, also a member of the Asteraceae, exhibited both fungicidal and anti-bacterial activities (Szakiel, 2005).

2.9 Saponins as Chemosystematic Markers

Until recently, the utilization of triterpenes and their glycosidic forms (the saponins) in chemosystematic studies have been limited. One recently published excellent review outlines the evolutionary development and distribution of these compounds in higher orders of angiosperms and summarizes findings on structural and distributional data obtained for triterpenoid saponins during the last 50 years (Henry, 2005). Based on these data, it is clear that triterpene-type saponins are restricted to more derived taxonomic groups such as the Carophyllideae, Rosideae and Asterideae, but are lacking in the Paleodicots and Monocots, where the steroidal saponins predominate. The only exception thus far was reported by Osbourn's group (Papadopoulou et al., 1999) that used a combination of phytochemical, molecular and genetic approaches to examine the expression patterns of naturally-occurring triterpene saponins found in oats (*Avena* spp), compounds that proved critical in defense against pathogenic fungi.

Recently, a review of the distribution and occurrence of saponins in the plant kingdom was published by Vincken et al. (2007); the aim of the review was to classify saponins based on the biosynthesis of their carbon skeletons and substitution patterns. Next the resulting classification was compared with the systematic position in plant

orders. The results from this analysis indicated that saponin biosynthesis does not appear to be plant-order specific, in regards to biosynthetic origin or with the attachments to the skeleton with various functional groups. Moreover, according to Vincken et al. (2007) only 18 species in the Asteraceae have been reported to contain saponins. This estimate did not include reference to *Silphium perfoliatum* saponins (Davidyants et al., 1984a,b; Davidyants et al., 1985; Davidyants et al., 1986) and a current literature search on Scifinder Scholar 2007 (the world's largest searchable collection of chemical substance information) retrieved over 100 references pertaining to "saponins" and "Asteraceae". Even when considering that saponins may have been reported from over a hundred species of Asteraceae, this still represents less than 1% of the entire family. The few literature reports on saponins from Asteraceae species may indicate that this family has not been extensively analyzed for these compounds and does not necessarily mean that saponins are not present throughout the family. The results outlined in this dissertation provide additional reports of new and previously described saponins from all *Silphium* species and from the closely related *Lindheimera texana*. These findings and other literature reports indicate that the Engelmanniinae subtribe, and perhaps many other unstudied Asteraceae species could be a rich source of new triterpene saponin structures.

Chapter 3: A Phylogenetic Analysis of Tribes of the Asteraceae Based on Phytochemical Data[1]

This chapter describes the first phylogenetic analysis of the entire Asteraceae based only on chemical data. The data matrix used in this study was based on a large chemical database comprising ~400 skeletal types of terpenes, coumarins, flavonoids, benzofurans and polyacetylenes isolated from the family. Hypotheses about the relationships among tribes are discussed based on the chemical data and compared with relationships inferred from the morphological and macromolecular based classifications.

3.1 Introduction

The secondary chemistry of the Asteraceae has been studied intensively over the last two centuries and several classification systems have been proposed based on combinations of chemical, morphological and, more recently, molecular data (Wagentiz, 1976; Carlquist, 1976; Seaman and Funk, 1983; Bremer, 1987; Karis et al., 1992; Bremer, 1994; Kim and Jansen, 1995; Bayer and Starr, 1998; Panero and Funk, 2002). Yet, no previous study has attempted to use only chemical data for a phylogenetic analysis of the entire Asteraceae. The widespread occurrence and trends of accumulation of certain classes of compounds make this family particularly well-suited for chemotaxonomic studies (Zdero and Bohlmann, 1990; Emerenciano et al., 1985; Emerenciano et al., 1986, Emerenciano et al., 1987; Alvarenga et al., 2001) and in some

[1] This chapter was previously published as: A Phylogenetic Analysis of Tribes of the Asteraceae Based on Phytochemical Data, by Lalita M. Calabria, Vicente P. Emerenciano, Marcelo J. P. Ferreira, Marcus T. Scotti and Tom J. Mabry, Natural Product Communications (2007) 1 (3) 277-285.

cases have proven useful in clarifying relationships among taxa at the tribal, generic and species levels (Alvarenga et al., 2001). In other cases, inconsistencies in the literature and great variability in chemical reports make chemical characters comparatively unreliable. Even today, little is known about the genetic and environmental controls over secondary metabolism and so presence/absence studies for the end products of biosynthetic pathways have often failed to predict the evolutionary relationships among taxa.

In recent years, some of these hurdles have been overcome by increasing the number of compounds isolated and then coupling this information with new programs for the manipulation of data and statistical analysis. These advances have allowed researchers to gain a much broader and more detailed view of the distribution and occurrences of secondary metabolites in the Asteraceae. In addition, the development of techniques to allow rapid DNA sequencing has provided innovative approaches to plant systematics and an opportunity to re-explore the usefulness of secondary metabolites for further refining molecular based systematic studies of the Asteraceae.

The familial classification of the Asteraceae began with the French botanist Henri Cassini, who in 1816 published a diagram showing the interrelationships of nineteen tribes of the Asteraceae (Figure 3.1) (Cassini, 1816). Then in 1873, Bentham revised Cassini's arrangement reducing the number of tribes to thirteen (Figure 3.2). The tribal subdivisions proposed by Bentham (Bentham, 1873) continue to be used even today, as part of the 11 subfamily and 35 tribes described in the recent molecular work of Panero and Funk (2000), whose molecular phylogeny included 13K of chloroplast sequence data and solidified the relationships among the tribes of Asteraceae.

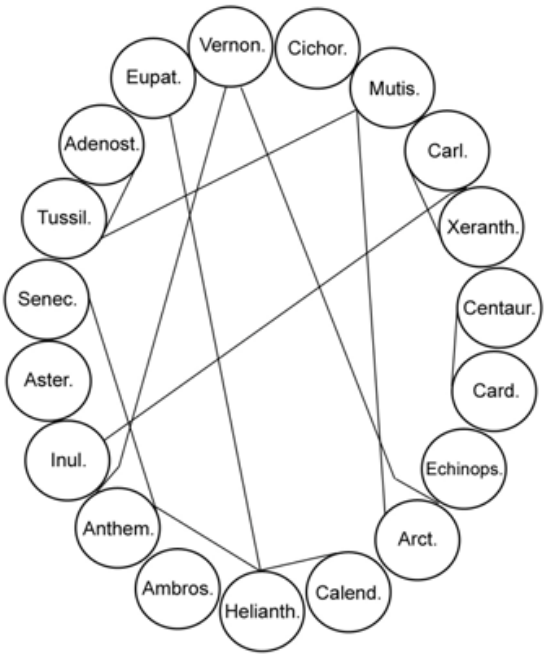


Figure 3.1. Diagram according to Cassini (1816) showing the interrelationships of nineteen tribes of the Asteraceae.

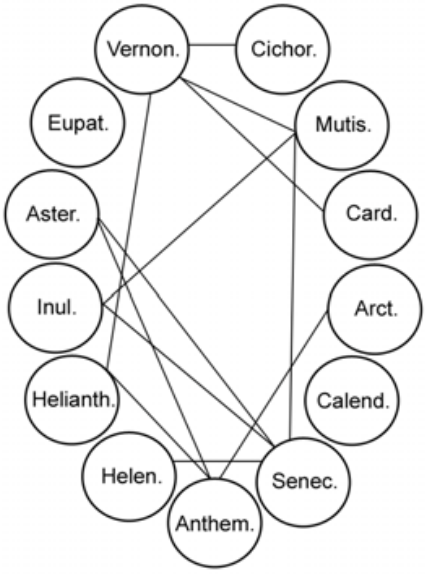


Figure 3.2. Diagram according Bentham (1873) reducing the number of tribes to thirteen from Cassini's arrangement (1816) for nineteen tribes.

Early work on the subfamilial classification of the family was carried out by several researchers. In 1976, Carlquist concluded that there are two subfamilies in the Asteraceae on the basis of morphological studies, the Asteroideae and the Cichorioideae (Table 3.1). In the same year Wagenitz also proposed a two subfamilial classification that differed from that of Carlquist by placing the Eupatorieae with the Asteroideae instead of in the Cichorioideae (Table 3.1). This biphyletic view of the family was a major step in understanding the tribal interrelationships within the Asteraceae (Carlquist, 1976; Wagentiz, 1976).

Wagenitz (1976)	Carlquist (1976)
I.	Subfamily Cichorioideae
Vernonieae	Mutisieae
Liabeae	Vernonieae
Mutisieae	Cardueae
Cardueae	Arctoteae
Echinopeae	Cichorieae
Arctoteae	Eupatorieae
II.	
Eupatorieae	Subfamily Asteroideae
Heliantheae	Heliantheae
Helenieae	Astereae
Senecioneae	Inuleae
Calenduleae	Calenduleae
Astereae	Senecioneae
Inuleae	Anthemideae
Anthemideae	

Table 3.1. Biphyletic classification systems of Carlquist (1976) and Wagenitz (1976) based on morphological characters.

In 1987, Bremer presented a working cladogram of the Asteraceae based on 81 characters, 10 of which were chemical (Figure 3.3) (Bremer, 1987). The remaining characters were mostly morphological features such as corolla type, styles and stamens and to a lesser extent chromosome numbers, anatomy and chloroplast DNA.

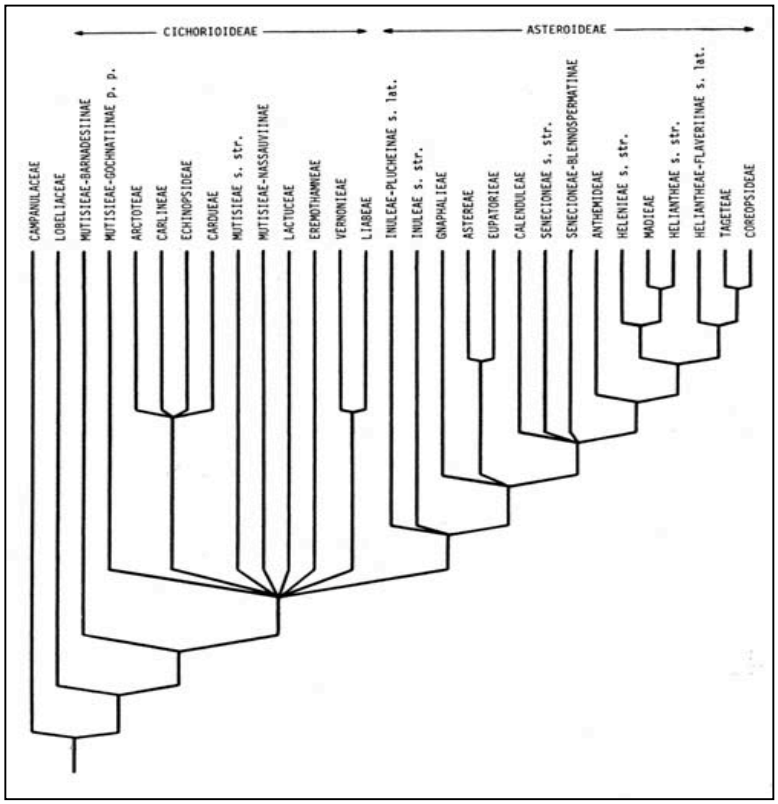


Figure 3.3. Bremer's (1987) strict consensus tree based morphological and chemical characters.

Bremer's study (1987) is one of the only examples of incorporating chemical characters in combination with morphological and molecular data into a cladistic analysis of the Asteraceae. Eight years later, at the Proceedings of the International Compositae conference held in London at Kew Gardens in 1995, Bremer presented a new subfamilial view of the Asteraceae based on a cladistic study of morphology in which he proposed

four subfamilies: the Asteroideae, Cichorioideae, Carduoideae and Barnadesioideae (Bremer, 1996). Bremer placed the Mutiseae as an unresolved clade nested between the Barnadesioideae and the Carduoideae. Figure 3.4 shows Bremer's diagram as presented in Kew (Bremer, 1996), which he modified from his original diagram published one year earlier (Bremer, 1994).

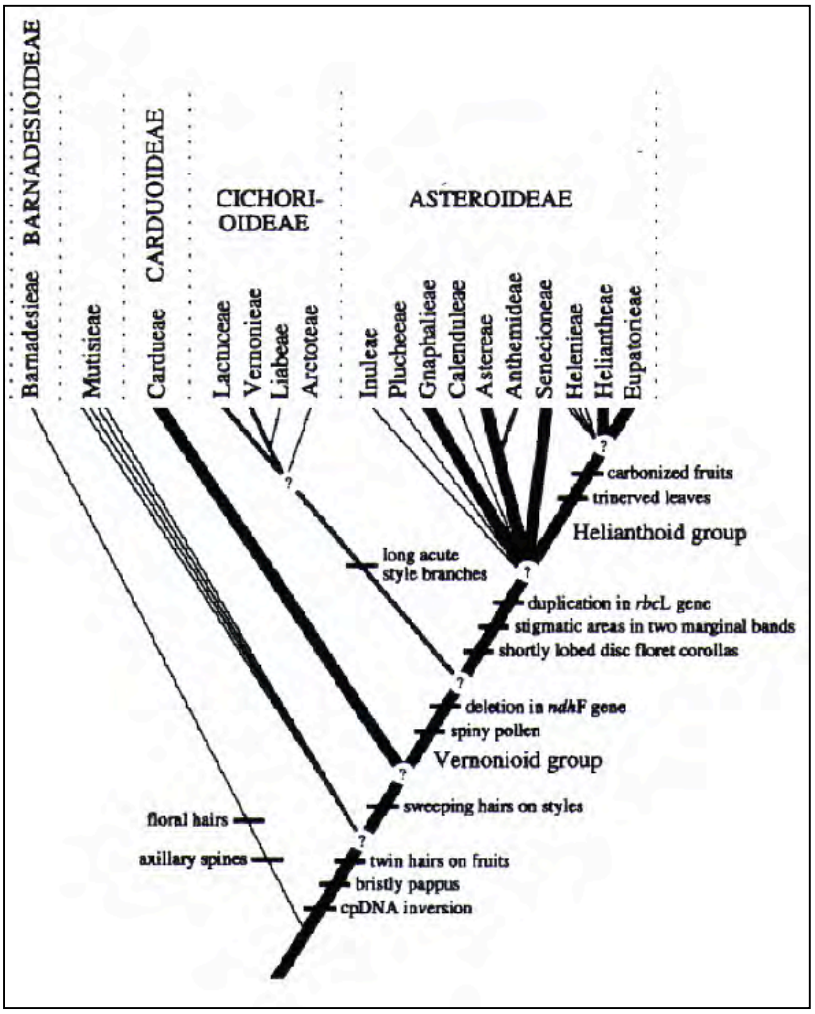


Figure 3.4. Morphologically-based diagram showing tribes of Asteraceae according to Bremer (1996). For the three-letter acronyms, see Table 3.2.

In 1995 at the Kew meeting, Kim and Jansen (1995, 1996) presented a phylogenetic analysis of 94 ndhF sequences representing all major clades of the Asteraceae (Figure 3.5). The ndhF chloroplast gene proved to be three times more phylogenetically informative than previously used genes, such as rbcL. The ndhF tree showed five major clades in the family and demonstrated the potential for this marker in resolving phylogenies.

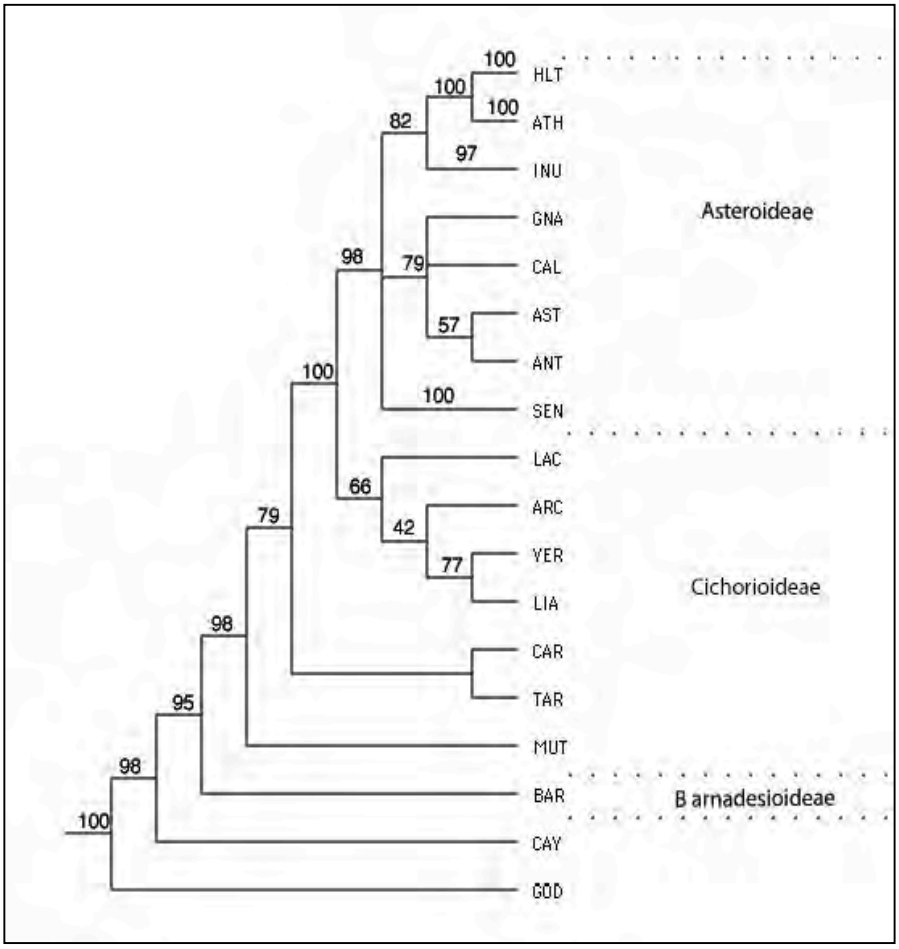


Figure 3.5. Kim and Jansen's (1995, 1996) strict consensus ndhF tree condensed to tribes. For the three-letter acronyms, see Table 3.3.

There are several ways in which the extensive chemical data that have become available can be used to understand the evolution and diversification of the Asteraceae. Perhaps the most useful application is the mapping of occurrences of different structural types onto a molecular phylogeny of the Asteraceae to visualize the distribution of chemical classes at different hierarchal levels. This approach was presented at the International Compositae Alliance meeting in Barcelona, Spain (July 2006) (Calabria et al., 2006) and will be published next year in the proceedings of the meeting. This application of the chemical data can provide clues to understanding how different chemical pathways have evolved in the Asteraceae.

Alternatively, chemical data can be applied to reconstructing phylogenetic trees in comparison with morphological or molecular-based phylogenetic analyses. Despite the limitations of using secondary metabolite data for phylogenetic reconstructions, there is still value in creating these chemical phylogenies in an attempt to understand why they differ from the molecular and morphologically trees. The results from these analyses should help to determine which clades of the Asteraceae are best defined by the chemical characters used in this study; the clades that exhibit a strong phylogenetic signal will be further examined to address questions regarding the importance of secondary metabolites in the adaptation and evolutionary success of the Asteraceae. In areas of the tree where chemical data produces arrangements of clades that differ from molecular and morphological-based studies, this information can also be useful in understanding where convergence leads to similar chemical expression patterns and can help answer questions

about the influence of secondary metabolites on the adaptive pressures and strategies used by the taxa.

3.2 Methods

In the chemically-based study presented here, genera of the Asteraceae were grouped into tribes following I. Bremer's morphologically-based classification (Bremer, 1994) and II. the molecular-based classification of Kim and Jansen based on ndhF sequence data (Kim and Jansen, 1995). It should be mentioned that the ndhF phylogeny of Kim and Jansen (1995), although one of the most important and influential papers related to Asteraceae phylogeny, is not the most up-to-date molecular phylogeny available. The recent molecular studies of Panero and Funk (2000) provide a more accurate phylogeny for the family, but do not provide a complete list of genera according to their new 35 tribal arrangement, and therefore, nearly half of the chemical data in our matrix would be missing if their tribal arrangement was employed for our analysis here. A phylogenetic analysis was carried out in PAUP (Swofford, 2002) using 389 chemical characters based on ~7,000 different compounds reported in the literature from 3,159 species of the Asteraceae, grouped by skeletal types (Figure 3.6) (Asfaw et al., 1999); Macleod and Rasmussen, 1999; Vichnewski et al., 1999; Sigstad et al., 1999; Petrovic et al., 1999; Gonzalez et al., 1997; Pari et al., 1998). The chemical characters used for the two outgroups, Goodeniaceae and Calyceraceae, were also taken from the literature [Ghisalberti, 2004; Bohm et al., 1995). The 389 skeletal types can be viewed at: http://www2.iq.usp.br/docente/vdpemere/skeltyp/index/.

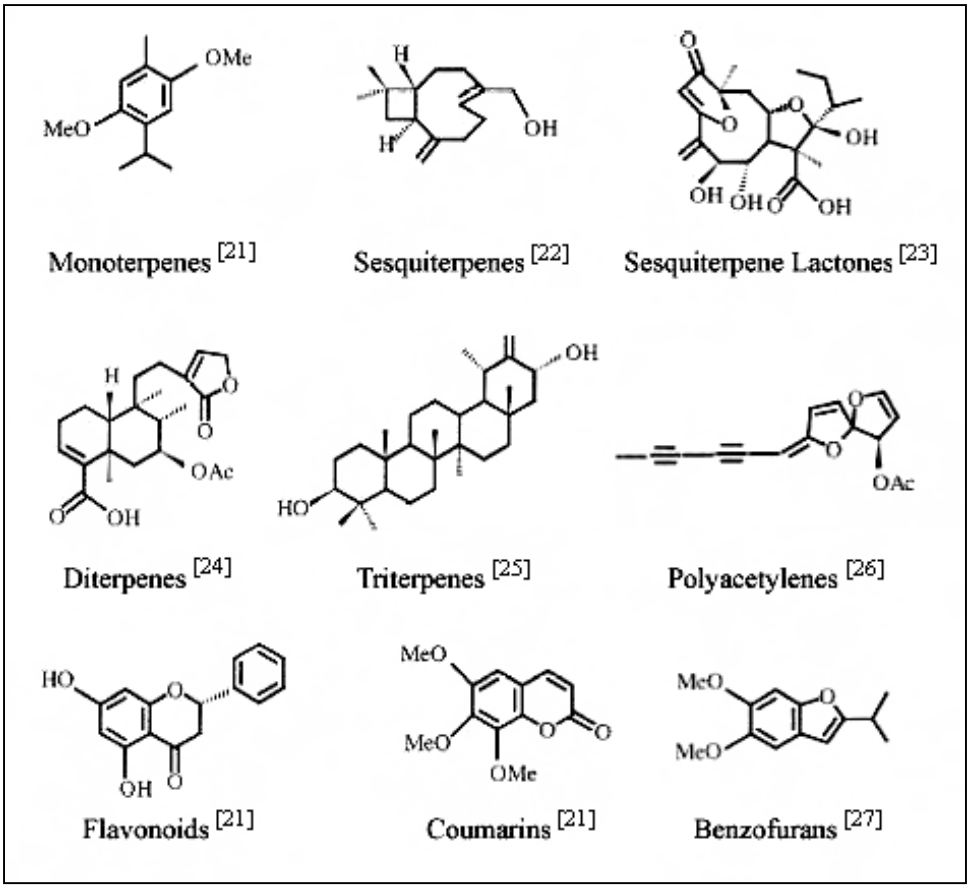


Figure 3.6: The 389 skeletal types used in this study fall into 9 chemical classes represented here by typical compounds found in the Asteraceae (see references listed next to each chemical class).

A data matrix was created for each tribal arrangement used in this study. The chemical data contained 3,159 species (666 genera) (Table 3.2) with a total of ~7,000 chemical compounds grouped into 389 skeletal types, thus compounds in each group are biogenically related. The data were extracted directly from the chemical database obtained from our own system, SISTEMAT (Gastmans et al., 1999), which includes the occurrence data (ca. 28,000 entries) for the 389 skeletal types examined in this study.

Taxon	Three-Letter Acronyms	No. of Species	No. of Species in the database
Anthemideae	ANT	1737	501
Arctoteae	ARC	139	26
Astereae	AST	2846	340
Athroisma group*	ATH	26	5
Barnadesieae	BAR	92	24
Calenduleae	CAL	113	28
Cardueae	CAR	2513	113
Calyceraceae*	CAY	50	4
Eupatorieae	EUP	2396	391
Gnaphalieae	GNA	1728	123
Gochnatieae	GOC	68	11
Goodeniaceae*	GOD	380	15
Helenieae	HEL	835	211
Heliantheae	HLT	2449	612
Inuleae	INU	480	174
Lactuceae	LAC	2486	115
Liabeae	LIA	159	17
Mutisieae	MUT	321	95
Nassauvieae	NAS	318	36
Plucheeae	PLU	220	26
Senecioneae	SEN	3247	196
Tageteae*	TAG	216	43
Tarchonantheae*	TAR	2	2
Vernonieae	VER	1346	131

Table 3.2: The taxa analyzed in this study and their respective three-letter acronyms, the number of species in each taxa and the number of species in the database. Nomenclature for tribes was adapted from Bremer (1994), except where there is an asterisk (*), which indicates nomenclature from Kim and Jansen (1995).

The assembly of the database arose through an extensive inspection of Chemical Abstracts from 1960 to 2004. The matrix used in this study consisted of taxa (subgroups of those listed in Table 3.2) scored for presence/absence of each of the 389 structural types. If a carbon skeleton appeared at least once in a given taxon then a value of one was assigned and if that skeleton was not present in the literature data then an assumed value of zero was given. Therefore, there were no missing data in the matrix.

Phylogenetic studies are based on the premise that characters and character states reflect common ancestry and thus will reconstruct the pattern of ancestry and descent. We have assumed here that the occurrence of the same chemical structural type in two taxa is the result of common ancestry although we realize that, like all characters, similar states can be the result of convergence.

Maximum parsimony (MP) analyses of the chemical data were performed using PAUP*4.0b10 (Swofford, 2002). Heuristic searches were performed using addition sequence with 1000 random additions of taxa, tree bisection-reconnection (TBR) branch swapping, and MulTrees (keeping multiple, shortest trees). A bootstrap analysis (Felsenstein, 1985) was also performed to evaluate the degree of support for branches using 1000 replicates with random sequence addition and TBR branch swapping. Bootstrap values of less then 50% are not shown. A strict consensus tree was produced from all equally parsimonious trees.

Genera from our chemical database of the Asteraceae (Gastmans, 1990) were assigned to tribes according to the morphologically-based classification of Bremer (1994) and entered into a data matrix with chemical characters. A maximum parsimony analysis of our data matrix organized by Bremer's tribal classification produced three shortest trees of 657 steps; 172 of the 389 chemical characters were parsimony informative. Figure 3.7 shows our strict consensus tree (I) with bootstrap values, the asterisks indicating groups not matching the classification system of Bremer (1994).

3.3 Results and Discussion

Our phylogeny based on chemical characters alone was able to differentiate the Barnadesioideae as a monophyletic group strongly supported (91% bootstrap) as sister to the rest of the Asteraceae family, which is in good agreement with Bremer's morphological findings (Bremer, 1994). The tribes placed by Bremer into the Cichorioideae were not grouped together by our chemically-based analysis but instead formed a grade between the Barnadesioideae and the Asteroideae. The position of the Carduoideae was unresolved in our chemically-based analysis, instead of being placed between the Cichorioideae and the Mutisieae as Bremer did in his analysis. The Mutisieae and the Vernonieae are grouped on the basis of chemical data as sister tribes for which there is moderate bootstrap support (70%). This grouping is probably due to the similarities in the production of complex coumarins by both tribes. It is interesting to speculate that perhaps the parallelism for the production of these compounds may have been a result of similar adaptive pressures. In contrast to the chemical data, Bremer placed Vernonieae as one of the four tribes known as the "Vernonoid group" comprising the Cichorioideae subfamily. Thus, the positioning of the Mutisieae and Vernonieae as sister to the Asteroideae subfamily based on the chemical data is unusual because the Mutisieae is almost always, after Barnadesioideae, sister to the rest of the Asteraceae. In our chemically-based analysis, the Liabeae and the Arctoteae of the Vernonoid group, both form monophyletic groups between the Barnadesioideae and the remaining tribes of the Asteraceae. Based on the chemical data in this analysis (Figure 3.7), the tribes that make up the Asteroideae subfamily formed a moderately supported (73% bootstrap)

monophyletic group, which agrees with morphological data (Bremer, 1994), with the exception of the Calenduleae and the Plucheeae, which are placed by the chemical data outside the Asteroideae. The chemical data only weakly supported intertribal relationships within the Asteroideae subfamily. This finding is similar to that of Bremer whose cladogram also showed a poorly resolved Asteroideae (Figure 3.4).

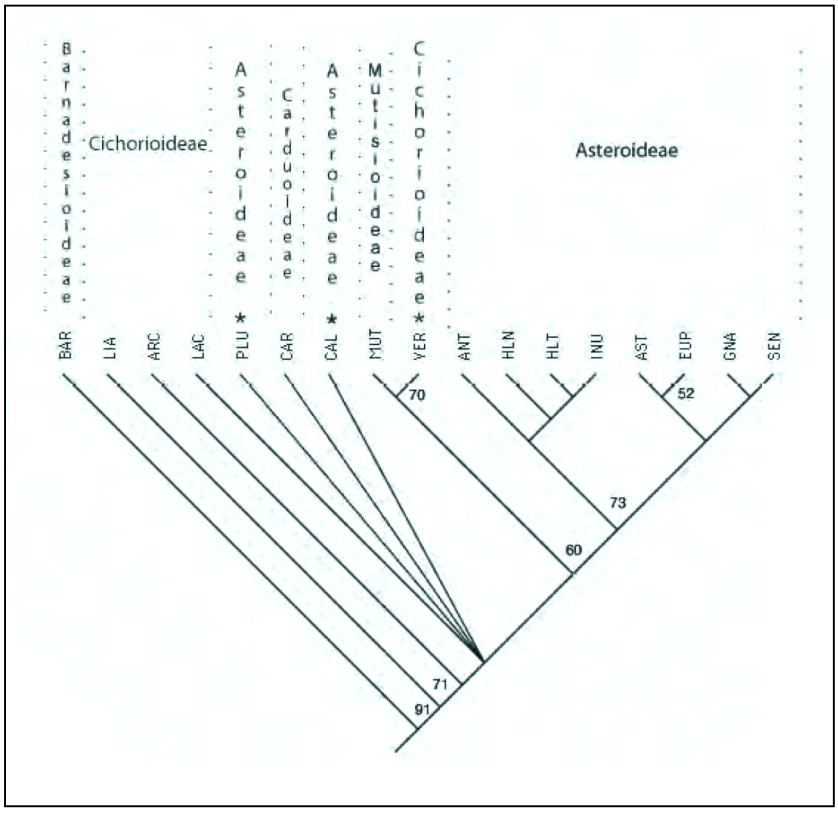


Figure 3.7: Our strict consensus tree (I) of three equally parsimonious trees with 657 steps based on chemical data using the tribal classification of Bremer (1994). The Goodeniaceae and Calyceraceae were used as outgroups in the parsimony analysis, but are omitted in this figure. For the three-letter acronyms, see Table 3.2.

Using a data matrix with genera placed in tribes according to Kim and Jansen's classification (Kim and Jansen, 1995) and scored for the 389 chemical characters, maximum parsimony produced three shortest trees each of 642 steps; 195 of the 389

chemical characters were parsimony informative. Figure 8 shows our strict consensus tree (II) with bootstrap percentages, the asterisks indicating tribes that differ in the subfamilial classification according to Kim and Jansen (1995, 1996). A major difference between our tree constructed from chemical data and the molecular based tree of Kim and Jansen (1995, 1996) is the polytomy between the outgroup Goodeniaceae and the ingroups Athroisma group and Tarchonantheae (Figure 3.8).

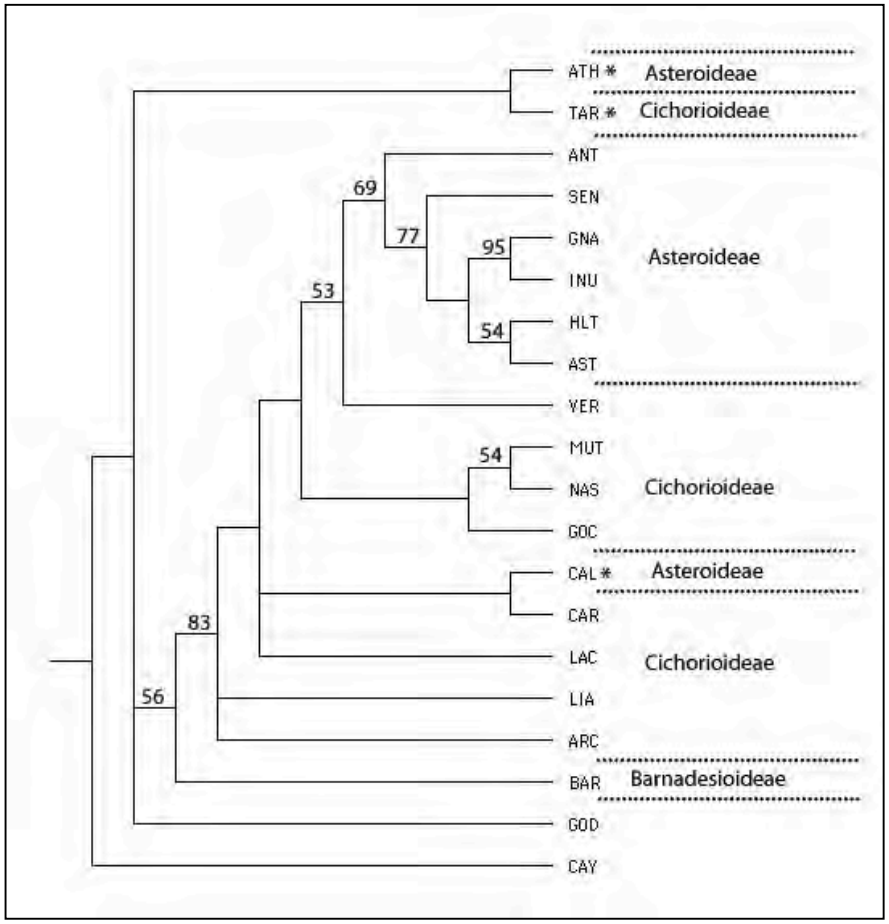


Figure 3.8: Our strict consensus tree (II) of three shortest trees of 642 steps based on chemical data using the tribal assignments of Kim and Jansen (1995). For the three-letter acronyms, see Table 2. Asterisks indicate taxa, which differ from Kim and Jansen's (1995) subfamilial placement.

The chemical data were unable to distinguish which of these two branches is sister to the Asteraceae. However, from molecular and morphological data, it is clear that both the Tarchonantheae and the Athroisma group are part of the Asteraceae family and that Goodeniaceae is an outgroup (Kim and Jansen, 1995; Jansen and Kim, 1996) A limited amount of chemical data and the small number of representative genera (one genus and two species representing ATH and one genus and one species representing TAR) can probably explain these differences in conclusions based on different types of data.

Our chemically-based phylogenetic analysis (Figure 3.8) showed a monophyletic Asteroideae with the exception of the Calenduleae, which instead was sister to the Cardueae. The bootstrap support for the Asteroideae branch was only 69% compared with 98% in Kim and Jansen's analysis and tribal affinities within the Asteraceae differed in several ways from Kim and Jansen's ndhF tree (1995, 1996). In Kim and Jansen's analysis the Asteroideae subfamily was a strongly supported monophyletic group, but within the subfamily the three monophyletic clades (i. Senecioneae; ii. the Gnaphalieae, Calenduleae Anthemideae and Astereae; iii. a clade which includes the Inuleae, as well as the Athroisma group and the Heliantheae *sensu lato*, which contains Helenieae, Coreopsideae, Eupatorieae and Tageteae) were not well resolved in their relationship to each other (Kim and Jansen, 1995; Jansen and Kim, 1996). Most tribes, however, were very well supported within each clade. In contrast, in our chemical analyses, intertribal relationships in the Asteroideae subfamily were fully resolved, although not all branches were well supported and some tribes were grouped differently from those based on Kim and Jansen's (1995) molecular data.

In our phylogenetic analysis based on chemical data using the tribal assignments of Kim and Jansen (1995), Anthemideae was sister to the rest of the Asteroideae, followed by Senecioneae, which in turn, was sister to two clades: one containing Heliantheae *sensu lato* and Astereae and another containing Inuleae and Gnaphalieae. Moreover, the group sister to Anthemideae was moderately supported with a bootstrap value of 77%, whereas the Astereae-Heliantheae *sensu lato* clade was weakly supported with a bootstrap value of 54%. The pairing of Gnaphalieae and Inuleae was the most strongly supported clade in our chemical analysis with a bootstrap value of 95%.

Paraphyly of Cichorioideae subfamily was indicated by the ndhF tree in Kim and Jansen's paper (1995). However, certain clades did form strongly supported monophyletic groups, such as the "LALV" group containing the Lactuceae, Arctoteae, Liabeae and Vernonieae, which had 100% bootstrap support as the sister group to the Asteroideae based on the ndhF data. In our chemical phylogenetic analysis the Vernonieae was the monophyletic sister group to the Asteroideae with 53% bootstrap support. It is worth mentioning that in previous chemical studies (Emerenciano et al., 2006) Vernonieae was found to be the group within the LALV clade that possessed the most advanced chemistry (based on its content of highly oxidized sesquiterpene lactones).

Our study also produced a paraphyletic Cichorioideae (Figure 8) with the tribes poorly resolved with the exception of the grouping of Nassauvieae and Mutisieae. Although only weakly supported by the chemical data in this study, this is the only case where the chemical data provided more bootstrap support than the molecular data (54% in this study compared with 50% (data not shown) in Kim and Jansen's (1995) molecular

based study). Because molecular and morphological data have both failed to resolve fully the relationships among members of this subfamily, this is an area of the tree that should be examined further using chemical data. There is a wealth of molecular and morphological data that supports the Barnadesioideae subfamily as sister to the rest of the Asteraceae (Bremer, 1994; Kim and Jansen, 1995; Panero and Funk, 2000). The phylogenetic analysis of the chemical data presented here moderately supports this view with a bootstrap value of 83%, with the exclusion of the unresolved polytomy containing the outgroup Calyceraceae and the Tarchonantheae and the Athroisma groups.

3.4 Conclusions

The purpose of this study was to compare previous classifications of the Asteraceae based on morphological (Bremer, 1994) and molecular data sets (Kim and Jansen, 1995) with new phylogenetic reconstructions based on chemical data of carbon skeletal classes compiled from extensive searches of the literature (strict consensus tree (I) and (II), Figures 3.7 and 3.8). The chemical characters used in this study allowed construction of tree topologies that differed in the relationship of some tribes from trees based on morphological and molecular data of the major clades of the Asteraceae. Not all of the chemical characters used in this study were informative and others have probably evolved independently in multiple lineages; thus the chemical data alone were unable to resolve relationships among some tribes. However, the chemical data defined the Barnadesioideae and the majority of the Asteroideae as distinct subfamilies. Relationships among some clades were moderately to well-supported by the chemical data and are in

good agreement with either the morphological or the molecular based classifications of the Asteraceae. Future work should focus on tribes and groups where the chemical data reflects the phylogenies produced with DNA data (i.e., Mutisieae and Nassauvieae). An in-depth chemical analysis of these clades may prove helpful in understanding the evolutionary role of secondary metabolites in these groups.

There are also cases where chemical data produce well supported but very different relationships among tribes or groups of taxa, when compared with relationships produced by molecular or morphological data. Many questions remain regarding why these differences in relationships occur between the chemical trees presented here and morphological and molecular based alignments. For example, why do clades that are strongly supported by chemical data in this analysis (i.e., GNA-INU) differ in their placement based on morphological and molecular data? There are, of course, many limitations of chemosystematics that need to be considered when interpreting the results of chemically-based studies (Van Wyk, 2003). Even with the growing number of reported taxa with secondary metabolite data appearing in the literature (28,000 chemical occurrences in the Asteraceae), there are still many genera and species for which no chemical data are reported. In addition, chemists spend a disproportionate effort on identifying new and unusual compounds that have potential medicinal and pharmaceutical value; moreover chemists looking for new structures will usually not isolate known compounds or compounds present in small quantities, making it even more difficult for chemosystematists to use chemical reports from the literature in their studies. The large quantitative and qualitative differences for different plant parts studied make it nearly

impossible to compare chemical data across the literature for a particular species, especially since the plant parts used in their studies are often not fully described.

Despite advances made both in phytochemistry and genetics, little is known about the genetic control of secondary metabolite production and this has a huge impact on the ability of scientists to use chemical data to infer evolutionary relationships, just as is often the case with morphological data. Synapomorphies and homplasy are present in both chemical and morphological data, but are difficult to distinguish. A plant's capability of producing a particular class of compounds is not necessarily reflected in the secondary metabolite patterns reported in the literature; thus the absence of a compound could mean the plant lacks the ability to produce this compound, but it could also mean the compound was not detected in that particular study. A deeper understanding of the genetic basis of secondary metabolism will be necessary for researchers to weight chemical characters in a way to gain the most informative results in phylogenetic analyses.

It is clear from the results presented here that the phylogenetic signal of the chemical data is compromised by the incomplete and often inconsistent chemical reports and until this problem with the literature data is overcome, molecular data will continue to provide the most reliable phylogenetic trees. However, a single view using one type of approach, whether it be molecular, morphological or chemical, cannot provide answers to all questions relating to the evolution of the Asteraceae. Therefore, there is still potential in using these data to answer important questions regarding the lineage of this family and the phylogenetic reconstructions based on the chemical data presented here provide a

starting point for addressing questions about the adaptive mechanisms and ecological functions of secondary metabolites in the Asteraceae.

Chapter 4: Secondary Chemistry of the Asteraceae[2]

The Asteraceae are one of the most chemically diverse groups of flowering plants and this has no doubt contributed to the success and diversification of this family. Almost 30 years after the publication of the first major work on the biology and chemistry of the Asteraceae, researchers continue to discover an astounding number of new compounds, and many species remain to be chemically analyzed. Here we outline our current understanding of the distribution, abundance and diversity of secondary metabolites in the Asteraceae in the context of a DNA-based phylogenetic tree, which is based on several recent molecular studies. The chemical classes that are covered include: flavonoids, sesquiterpene lactones, mono-, di- and triterpenes, polyacetylenes, alkaloids, coumarins and benzofurans. We summarize the information for secondary metabolites in this family from extensive databases established in recent years, as well as available data in the literature and from communications with colleagues.

[2] This chapter will be published as part of the proceedings of the International Compositae Alliance on the Evolution and Systematics of the Asteraceae. The book is currently in press as: "The Secondary Chemistry of the Asteraceae; by Lalita M. Calabria, Vicente P. Emerenciano, Marcus T. Scotti, and Tom J. Mabry

4.1 Introduction

The purpose of this chapter is to outline the most current data available on the secondary chemistry of the Asteraceae in the context of a DNA-based phylogenetic framework. We describe the importance of chemical data for understanding evolutionary and systematic aspects of the Asteraceae as we review several examples at different hierarchal levels. We suggest that phytochemical and DNA-based phylogenetic data can complement each other, providing insights into the biological and ecological functions of secondary metabolites produced by this family. Moreover, secondary metabolite data can be used to solve difficult taxonomic problems in situations where DNA or morphological data alone cannot fully resolve relationships among taxa.

4.2 Background and Historical Perspectives

The great morphological and geographical diversity of the Asteraceae is reflected in its ability to produce a wide range of secondary metabolites including: monoterpenes, diterpenes, triterpenes, sesquiterpenes and sesquiterpene lactones, polyacetylenes, flavonoids, phenolic acids, benzofurans, coumarins and pyrrolizidine alkaloids (these alkaloids are, with few exceptions, confined to the Senecioneae and the Eupatorieae). The main biosynthetic pathways are derived from acetyl coenzyme A and/or amino acids leading to a remarkable number of unique structures isolated from nearly every tribe and genus investigated. Figure 4.1 shows an example of each of the main chemical classes found in the Asteraceae.

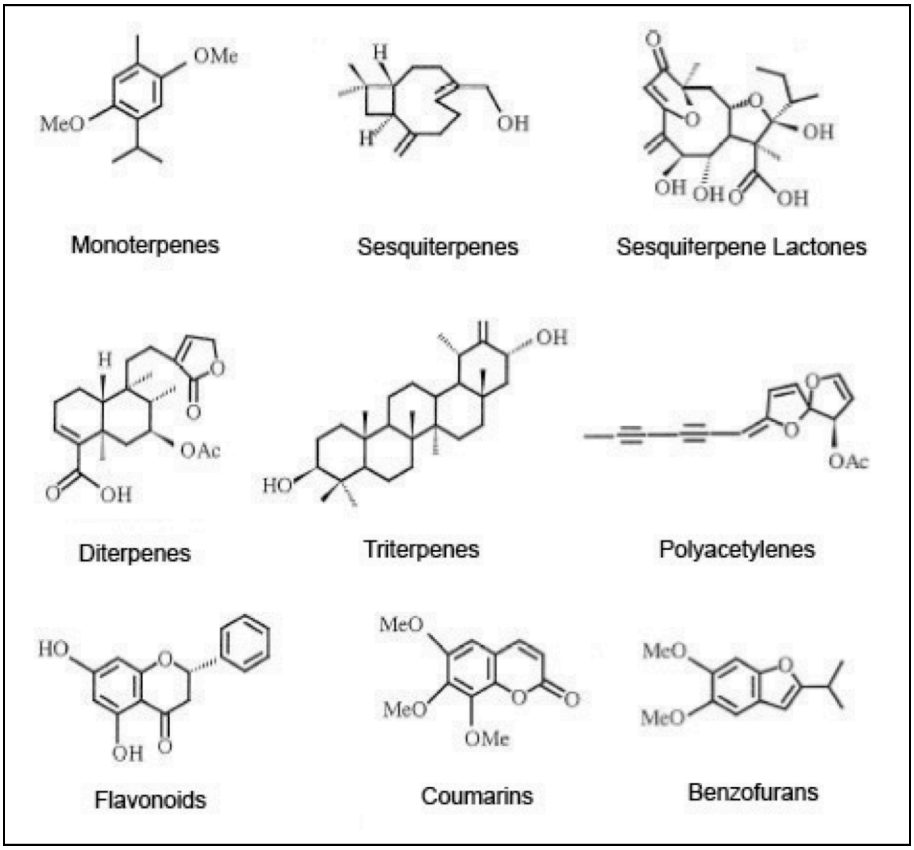


Figure 4.1. An example of the nine major chemical classes in our chemical database represented by typical compounds found in the Asteraceae. References for each compound (from left to right, top to bottom): monoterpenes (Asfaw et al., 1999), sesquiterpenes (Macleod and Rasmussen, 1999), sesquiterpene lactones (Vichnewski et al., 1999), diterpenes (Sigstad et al., 1999), triterpenes (Petrovic et al., 1999), polyacetylenes (Gonzalez et al., 1997), flavonoids (Asfaw et al., 1999), coumarins (Asfaw et al., 1999), benzofurans (Pari et al., 1998).

Due to the large number of species that exhibit diverse chemical profiles it is not surprising that the secondary chemistry of the Asteraceae has always been an area of intense interest to plant chemists and systematists. Several researchers have made important contributions to the field of chemosystematics, that is, the application of chemical data for evolutionary studies in plants (eg. Mabry, 1973; Harborne, 1977; Waterman, 1987; Gottlieb, 1982; Hegnauer, 1996 and all of their co-workers), many of

them devoting significant portions of their careers to chemical studies of this family. However, the once dynamic field of chemosystematics has been essentially displaced in the last 30 years with increasing focus on molecular-based DNA studies. The development of technology to allow rapid DNA sequencing has provided innovative approaches to plant systematics and an opportunity to re-examine phytochemical data in the context of well-resolved phylogenies for the Asteraceae. The structural characterization of secondary metabolites is a necessary part of any plant chemistry investigation and with close to a century of chemical data collected for this family, researchers now have a much broader and more detailed view of the distribution and occurrences of secondary metabolites for the family.

One of the first major syntheses on the chemistry of the Asteraceae was published in the 1977 volumes entitled "The Biology and Chemistry of the Compositae". In the final chapter, Mabry and Bohlmann (1977) state that "probably every member of the family contains flavonoids and most contain volatile oils and triterpenes." They also recognized the, "….presence of two unique structural classes of compounds which characterize the family; sesquiterpene lactones and acetylenes" and agreed that "….no major classes of alkaloids have been reported with the exception of pyrrolizidine alkaloids in the Senecioneae and, to a lesser extent, in the Eupatorieae." The trends observed for the distribution and accumulation of these compounds were recognized early in studies of the Asteraceae and today are supported by an even larger body of chemical data.

When Mabry and Bohlmann's review (1977) was published, the Barnadesieae were considered to be one of the more derived lineages in the Asteraceae whereas the Heliantheae were considered a basal lineage. At the end of their chapter they summarized data for two alternative hypotheses that addressed the question "are chemically complex tribes primitive or is it chemically depauperate tribes that are primitive?" If the former concept is correct then early diverged members of the Asteraceae would possess a wide range of secondary metabolites and advancement would be indicated by successive losses of compounds by more derived groups. If the latter were true then more "primitive" groups would be characterized by a simple chemical profile, while advanced groups would be represented by chemical differentiation and gains of new secondary metabolites. Indeed this topic sparked much debate for years after the 1977 paper was published!

In the late 1980's, the discovery of a chloroplast genome inversion shared by all members of the Asteraceae except the subtribe Barnadesiinae confirmed its collinear position with the rest of flowering plants and clearly established its place as the primitive group within the Asteraceae (Jansen and Palmer, 1987). In light of the many advances made in both phytochemical and taxonomic methods since then, scientists now understand that chemical pathways have differentiated through biosynthetic simplification and/or diversification at all taxonomic levels of the Asteraceae.

The next major work on the chemistry of the Asteraceae was published by Zdero and Bohlmann (1990) in an article summarizing important milestones in our understanding of the chemical diversity found in the family. They outlined over 7,000

constituents identified from over 5,000 species studied before 1990 and provided a broad assessment of the available chemical data. The authors pointed to the special trend of accumulation of lactonized sesquiterpenes and polyacetylenes, as well as the occurrence of many highly oxidized compounds, a pattern that is now considered the signature chemical profile shared by the majority of the Asteraceae.

The contributions of Bohlmann and co-workers to the field of Asteraceae chemistry continues today with an excellent online resource known as “The Bohlmann Files,” which provides a searchable database for secondary metabolites found in the Asteraceae (Berendsohn et al., 1998). The database was compiled from the original card catalogue created by Bohlmann and co-workers to document the thousands of secondary metabolites isolated and characterized by members of his research group.

Coinciding with the intensive chemical studies of Asteraceae in the last half of the 20th century was the development of a thriving chemosystematics program directed by the late Otto Gottlieb at the University of São Paulo in Brazil. Gottlieb believed that plant chemist’s should focus their efforts on examining similarities in biosynthetic pathways, not the substances produced, as an indication of phylogenetic relatedness (Gottleib, 1982). His work covered terpenoids, alkaloids, coumarins, xanthones, phenylpropanoids, polyacetylenes, iridoids among other secondary metabolites. Gottlieb’s school marked the first major attempt to measure chemical evolution using phenetic methods for groups of Angiosperm taxa.

The methodology developed by Gottlieb and collaborators was greatly influenced by the classic work "Numerical Taxonomy" of Sneath and Sokal (1973), whose studies

were based on cluster analyses using similarity matrices. Numerical taxonomy measures the gradual change in oxidation states of "micromolecules" (secondary metabolites), indicating the tendencies of the evolutionary processes that can be expressed as numeric indexes (Richardson and Young, 1982; Stuessy and Crawford, 1983; Emerenciano et al., 1987; Levy, 1997). Another important measure of the evolutionary process in plants was calculated using probable biogenetic maps for each chemical class by observing the number of carbon-carbon connections formed or broken in a biogenetical sequence. Although Gottlieb's evolutionary approach was more phenetic than phyletic, homology of biosynthetic routes for the establishment of phylogenies would later be applied to solving taxonomic problems in the Asteraceae.

Then in 1985, a doctoral student in Gottlieb's research group, Vicente Emerenciano, introduced the application of computational techniques for the organization and analysis of the chemical data (Emerenciano et al., 1985, 1986, 1987). In the last two decades the research group at the Institute of Chemistry of the University of São Paulo created a system for chemotaxonomy and structural determination of secondary metabolites found in the Asteraceae. This system, called "SISTEMAT," can codify the structure of a natural product and associate the carbon skeleton, not only to the botanical origin, but also to corresponding spectral data (Gastmans et al, 1990a and 1990b).

Using the chemical database SISTEMAT, researchers at the University of São Paulo continued their studies of the Asteraceae, applying a broad range of statistical methods to establish phenetic relationships among members of the Asteraceae. Earlier studies using SISTEMAT employed techniques such as PCA (Principal Component

Analysis) implemented in commercially available programs (Statistica©, Unscrambler®) to examine several metabolites simultaneously, for example, in PCA analysis correlating the presence of sesquiterpene lactones and flavonoids (Emerenciano et al., 1987). Later, Emerenciano and collaborators published a PCA analysis for tribes of Asteraceae using all major classes of secondary metabolites found in the Asteraceae (Alvarenga et al., 2001). The SISTEMAT database has been employed for several other applications including, testing of the redox theory in plants (Gottlieb, 1993) using PLS (Partial Least Squares) analysis (Emerenciano et al., 2006), PCA analysis of Heliantheae, and for the analysis of flavonoids hydroxylation patterns in tribes of the Asteraceae (Emerenciano, 2001).

Most recently, members of Emerenciano's group in collaboration with researchers at the University of Texas at Austin, published the first phylogenetic analysis for the entire Asteraceae based only on phytochemical data (Calabria et al., 2007). A cladistic analysis employing sesquiterpene lactones as chemical characters was previously published by Seaman and Funk (1983), however, the authors discussed only two examples from the Asteraceae. The data matrix used by Calabria et al. (2007) was based on a large chemical database comprising ~400 skeletal types of secondary metabolites isolated from the family. Previous classifications based on morphological and molecular data sets were compared with new phylogenetic reconstructions based on chemical data providing a novel context for addressing questions regarding the evolution of secondary metabolism in the Asteraceae.

4.3 Methods

The chemical data were compiled by extensive inspection of Chemical Abstracts from 1960 to 2006 and were entered directly into an Excel file containing approximately 3,024 species from 546 genera of Asteraceae, with a total of ~ 10,000 different chemical compounds grouped into nine major chemical classes (Figure 4.1). Throughout this chapter the term, "chemical occurrence" will be used to describe the number of times a compound of a given chemical class was reported in the literature for each species, genera, tribe, etc. For example, if two sesquiterpene lactones of the same skeletal type were isolated from *Silphium albiflorum* and one of these sesquiterpene lactones was also isolated from *Silphium perfoliatum*, then the total chemical occurrences of sesquiterpene lactones for *Silphium* would equal three, representing two different compounds, from two different species, one genus and one tribe.

Genera of the Asteraceae were grouped into tribes according to the recent molecular phylogeny constructed by Panero and Funk (2002). However, we included the Gundelieae as a separate tribe and the Plucheeae was placed as a monophyletic clade nested within the Inuleae tribe following the supertree of Funk et al. (2005), for which the Panero and Funk (2002) phylogeny served as a backbone for combining individual phylogenies for each tribe. Assignments of genera into tribes that were not outlined by Panero and Funk (2002) or Funk et al. (2005), were taken from Panero (2007). In addition, for subtribal, generic and species level analyses, we assigned genera into tribes according to Clevinger and Panero (2000). Table 4.1 shows the tribal names with the 3-letter codes used in this study and the approximate number of genera and species in each

tribe compared with the number of genera and species represented in the chemical matrix. The subfamilies Hecastocleidoideae and Gymnarrhenoideae (both monotypic taxa) had no chemical reports in the literature.

Tribe	Three-Letter Acronym	No. of Genera	No. of Genera in database	No. of Species	No. of Species in database
Anthemideae	ANT	111	28	1800	493
Arctotideae	ARC	17	11	215	38
Astereae	AST	205	60	3080	338
Athroismeae	ATH	6	3	59	6
Bahieae	BAH	20	8	83	23
Barnadesieae	BAR	9	8	91	27
Calenduleae	CAL	12	7	120	35
Cynareae	CYN	73	28	2360	240
Chaenactideae	CHA	3	1	29	2
Cichorieae	CIC	86	23	1500	52
Coreopsideae	COR	30	10	550	137
Corymbieae	CRY	1	1	9	1
Dicomeae	DIC	7	2	100	6
Eupatorieae	EUP	182	75	2200	333
Gnaphalieae	GNA	185	24	1240	178
Gochnatieae	GOC	4	5	62	16
Gundelieae	GUN	2	2	2	2
Helenieae	HLN	13	11	120	66
Heliantheae	HLT	113	73	1500	352
Inuleae	INU	66	27	687	92
Liabeae	LIA	16	6	190	10
Madieae	MAD	36	14	200	92
Millerieae	MIL	34	24	400	80
Mutisieae	MUT	64	26	715	44
Neurolaeneae	NEU	5	5	150	16
Perityleae	PER	7	5	84	13
Pertyeae	PRY	4	4	70	2
Polymnieae	POL	1	1	3	2
Senecioneae	SEN	150	25	3500	174
Tageteae	TAG	32	9	270	30
Tarchonantheae	TAR	2	1	15	9
Vernonieae	VER	118	21	1000	117

Table 4.1. Tribes of the Asteraceae analyzed in this study and their respective three-letter acronyms, the number of species in each tribe and the number of species recorded in our chemical database. Nomenclature for tribes follows Panero and Funk (2002), Funk et al. (2005) and Jeffery (2007).

Table 4.2 shows the data matrix created with the tribal arrangements used in this study and the number of occurrences of each class of secondary metabolites isolated from each tribe of the Asteraceae. For figures with chemical occurrence data mapped on phylogenies for the Asteraceae, the data matrix shown in Table 4.2 was converted into "one's" and "zero's" representing the "presence" or "absence" of a particular class of secondary metabolites in each taxon. "Presence" and "absence" refer to whether or not a

given chemical class was reported in the literature data and, ultimately, recorded in our chemical database for the Asteraceae.

4.4 Results and Discussion

Here we outline our present knowledge of the distribution, abundance and diversity of secondary metabolites in the Asteraceae and discuss useful approaches for detecting chemical patterns at the subfamilial, tribal, subtribal, generic and species levels. Figure 4.2 shows the total chemical occurrences for each main class of secondary metabolites found in our chemical database for the Asteraceae.

TRIBE	DITE	LACT	TRIT	SESQ	MONO	COUM	BENZ	POLY	FLAV	Totals
BAR	0	0	8	0	0	4	0	12	16	**40**
MUT	9	15	27	126	30	151	6	340	653	**1357**
GOC	20	7	8	4	0	3	0	15	22	**79**
DIC	15	26	16	1	0	0	0	17	18	**93**
CYN	130	322	23	87	22	30	1	163	303	**1081**
TAR	17	24	5	6	28	0	0	39	73	**192**
PRY	0	4	1	0	0	0	0	1	1	**7**
GUN	0	6	5	1	0	3	0	0	1	**16**
CIC	28	163	19	36	15	53	2	125	231	**672**
ARC	28	24	35	1	1	1	0	38	41	**169**
VER	288	214	174	10	24	23	1	232	290	**1256**
LIA	15	37	9	8	5	0	0	22	35	**131**
CRY	10	0	0	0	0	0	0	0	0	**10**
SEN	60	111	37	837	130	24	108	1136	2235	**4678**
CAL	94	0	11	18	7	3	7	46	81	**267**
GNA	301	3	29	64	145	46	19	303	577	**1487**
ANT	309	810	28	236	3454	212	4	3934	7840	**16827**
AST	971	19	63	99	626	173	55	1016	1969	**4991**
INU	122	242	15	271	330	36	3	655	1295	**2969**
ATH	4	21	0	0	3	0	0	3	6	**37**
HLN	4	359	0	4	78	1	0	83	166	**695**
COR	5	5	1	19	13	4	2	39	77	**165**
NEU	40	63	1	18	16	0	0	35	69	**242**
TAG	4	0	0	19	212	7	4	242	484	**972**
CHA	2	8	0	0	0	0	0	0	0	**10**
BAH	78	53	5	0	32	4	0	41	77	**290**
POL	16	0	0	0	0	0	1	1	2	**20**
HLT	883	771	45	122	302	32	177	678	1311	**4321**
MIL	528	154	8	28	91	2	2	131	254	**1198**
MAD	22	16	2	8	44	9	5	68	134	**308**
PER	2	1	0	0	7	0	0	7	14	**31**
EUP	885	370	85	151	389	37	208	870	1655	**4650**
Totals	**4890**	**3848**	**660**	**2174**	**6004**	**858**	**605**	**10292**	**19930**	**49261**

Table 4.2. Raw data matrix created for this study showing the number of occurrences of each class of secondary metabolites recorded in our chemical database from each tribe of the Asteraceae.

It is not surprising that flavonoids are nearly twice as abundant (~20,000 chemical occurrences) as any other chemical class evaluated in this study. Flavonoids serve diverse functions in plants and the subject of their ecological significance has been extensively reviewed (Bohm, 1998; Seigler, 1998; Bohm and Stuessy, 2001). Flavonoids attract pollinators (yellow flower color predominates in the Asteraceae) and seed and fruit dispersers. They provide protection against UV light and function in numerous plant-plant and plant-microbe signaling interactions. Bohm and Stuessy (2001) provide an excellent review of the distribution and occurrence of flavonoids in the Asteraceae, outlining the main structural types found at different hierarchal levels: anthocyanins, chalcones, aurones, flavanones, flavones, and flavonols. Similar to their review, we found flavones and flavonols to be the most commonly occurring flavonoid structural types in our chemical database for the Asteraceae.

The polyacetylenes occur over 10,000 times in our chemical database (Figure 4.2), and although their distribution is not as consistently widespread as the flavonoids, they do tend to accumulate in large amounts in certain taxa. Four separate reviews were published (Christensen and Lam, 1990; Christensen, 1991a, 1991b, 1992) outlining the distribution of polyacetylenes in the tribes Cynareae, Senecioneae, Astereae, and Anthemideae. According to our chemical database polyacetylenes reach their maximum abundance in the Anthemideae with nearly 4,000 chemical occurrences, followed by the Heliantheae and the Astereae, (both ~1000 chemical occurrences, Table 4.2). The Eupatorieae, Heliantheae and Inuleae also accumulate significant amounts of polyacetylenes according to our chemical database. Polyacetylenes serve as important

defense compounds in the Asteraceae, exhibiting insecticidal and anti-feedant activities (Seigler, 1998). Some polyacetylenes are considered phytoalexins, their concentration rapidly increases in response to attack by pathogenic fungi (Seigler, 1998). In addition, polyacetylenes found in the Asteraceae display potent phototoxic effects, killing or inactivating viruses, bacteria, fungi, nematodes and also negatively affecting herbivorous insects.

Terpenoids are the largest class of secondary metabolites found in plants, including monoterpenes (C_{10}), sesquiterpenes (C_{15}), diterpenes (C_{20}) and triterpenes (C_{30}) and are the second most common metabolites in our database for the Asteraceae. Many of these compounds serve as volatile signals in plant-insect and plant-plant interactions and play essential roles in plant reproduction and defense. Sesquiterpenes (including lactones) and monoterpenes are equally abundant in the Asteraceae with ~6,000 chemical occurrences recorded in our database (Figure 4.2). Diterpenes are also relatively abundant in the Asteraceae, occurring ~4,800 times in our chemical database (Figure 4.2). Alvarenga et al. (2005) recently published a detailed review of the chemosystematic importance of diterpenes in the Asteraceae. In addition, an extensive overview of the distribution and occurrence of diterpenes was given by Seaman et al. (1990).

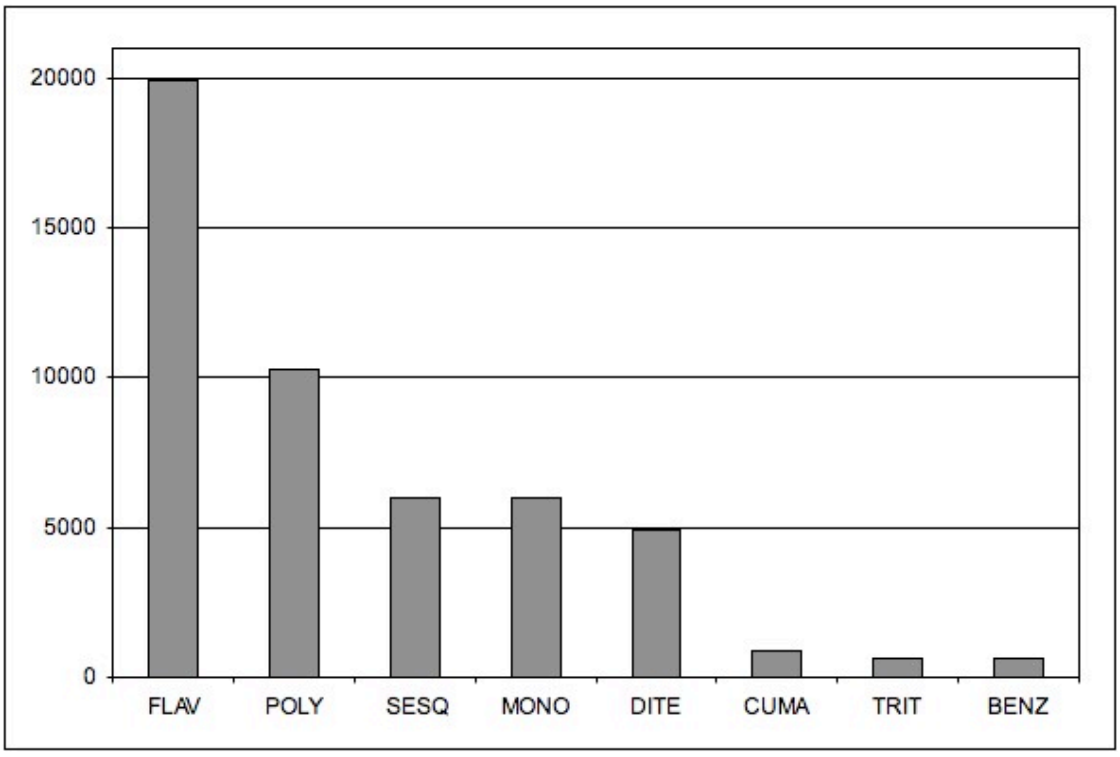


Figure 4.2. Total chemical occurrences for the Asteraceae representing each class of secondary metabolites in our chemical database.

The remaining classes of secondary metabolites are present in relatively low abundance in our chemical database for the Asteraceae, with fewer than 1,000 chemical occurrences (Figure 4.2). Although triterpenes are one of the least abundant chemical classes they occur in 28 of the 35 tribes of the family. This pattern of occurrence is consistent with their biological significance because triterpenes belong to the same biosynthetic group as steroid hormones and exhibit their physiological effects at very small doses. Saponins, which are the glycosidic forms of triterpenes, play important roles in the structure and function of cell membranes in addition to their role as defensive compounds against insects and pathogens.

Coumarins are well distributed in the Asteraceae, but do not occur consistently in all tribes or subfamilies. The only tribes that accumulate coumarins (over 150 occurrences) are the Anthemideae, Astereae and Mutisieae (Table 4.2). A group of complex coumarins known as furanocoumarins are common in the family and display a variety of biological activities against insects and pathogens. Like polyacetylenes, furanocoumarins also possess potent phototoxic properties capable of killing or inhibiting growth of pathogens, insects and nematodes (Seigler, 1998). Benzofurans occur sporadically throughout the Asteraceae and, except for triterpenes, are the least abundant chemical class represented in our database. Figure 4.3 shows the distribution of benzofurans in tribes of the Asteraceae according to Panero and Funk (2002) and Funk et al. (2005). The Eupatorieae, Heliantheae and Senecioneae are the only three tribes that have more than 100 chemical occurrences of benzofurans according to our database.

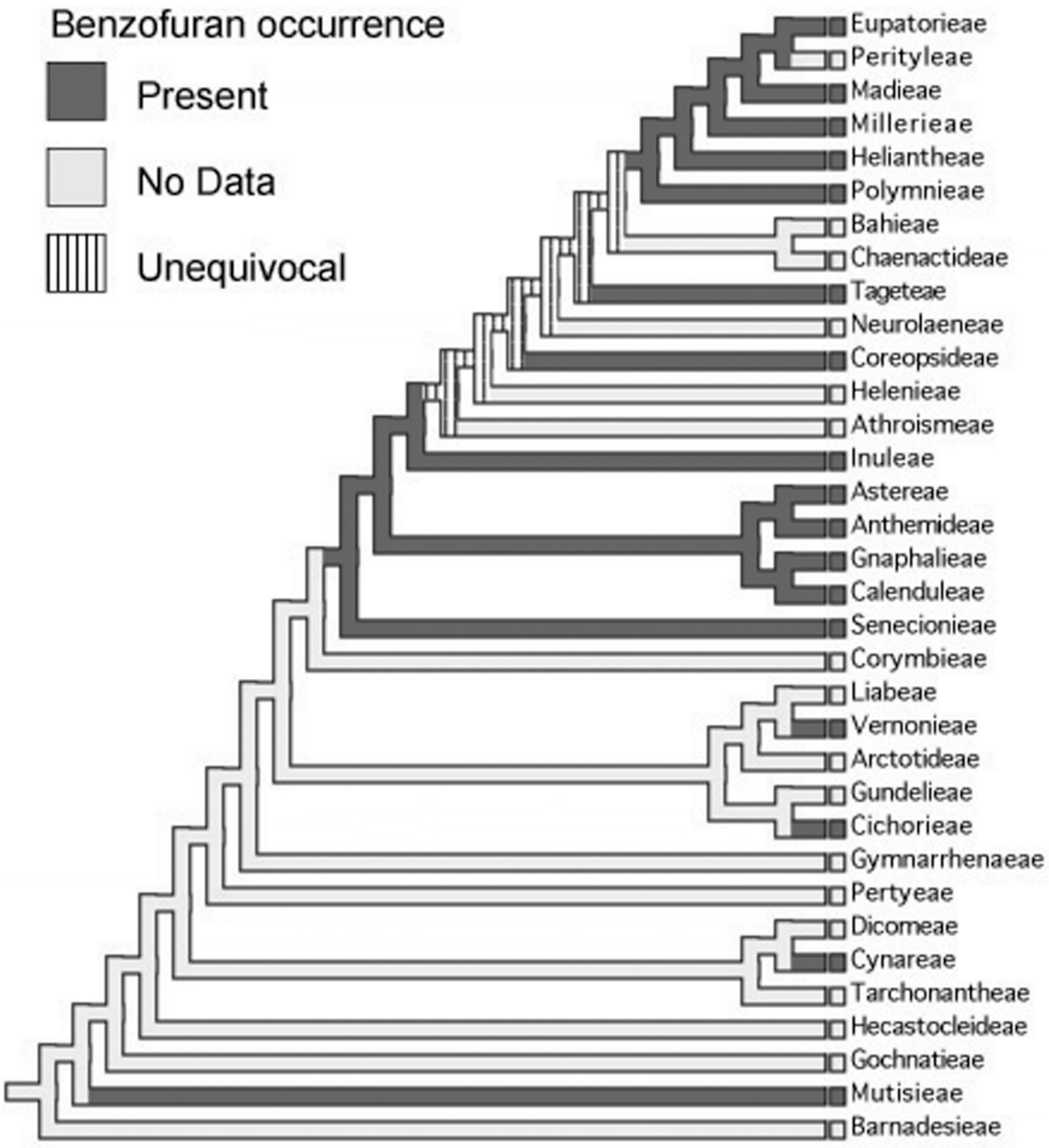


Figure 4.3. Distribution of benzofurans viewed at the tribal level, expressed as "presence" or "absence" of chemical data recorded in our chemical database for the Asteraceae. DNA-based phylogenetic framework adapted from Panero and Funk (2002) and Funk et al. (2005).

The distributional trend observed in Figure 4.3 suggests that benzofurans occur more consistently in the derived lineages of the Asteraceae. Although, it is unclear whether the chemical occurrence patterns observed for benzofurans represent their natural distribution or whether these compounds are investigated less frequently than other chemical classes in the family. For example, reports of triterpenes in the subtribe Engelmanniinae (Heliantheae) were limited to four genera prior to our chemical investigations, which found that every genus in this subtribe contains triterpene glycosides (data unpublished). Proksch and Rodgriguez (1983) reviewed the distribution and biological significance of benzofurans and chromenes (benzopyrans) in the Asteraceae; however, no other major review of these compounds has been published.

The pyrrolizidine alkaloids are not represented in our chemical database for the Asteraceae because of their limited occurrence. Figure 4.4 shows the distribution of pyrrolizidine alkaloids on a phylogenetic tree based on DNA-data (Panero and Funk, 2002; Funk et al. 2005). Historically, pyrrolizidine alkaloids were thought to accumulate only in the Eupatorieae and the Senecioneae although these tribes are not phylogenetically closely related. There is evidence to support the polyphyletic origin of pyrrolizidine alkaloids in the Asteraceae based on differential tissue expression of homospermidine synthase (HSS), the main enzyme involved in the biosynthesis of pyrrolizidine alkaloids (Anke et al., 2004). The authors concluded that within the Senecioneae and Eupatorieae HSS is expressed in different tissue types and therefore, the ability to produce these compounds arose multiple times independently.

Since the original volumes of the Biology and Chemistry of the Compositae (1977) were published, alkaloids have been reported from several tribes in addition to the Senecioneae and the Eupatorieae; the Cynareae *(Cirsium, Carduus, Echinops and Centaurea*) (Jordon-Thaden, 2003; Chaudhuri, 1992; Hymete et al. 2005; Sarker et al, 2001), Madieae (*Arnica, Melampodium)* (Passreiter, 1992; Schüngel and Passreiter, 2000), Heliantheae (*Echinacea*) (Roeder et al, 1984) and Neurolaeneae (*Neurolaenea*) (Passreiter, 1998). However the pyrrolizidine alkaloids in the latter three tribes are not considered "true" alkaloids, but rather β-amino acids, due to their atypical structures and biosynthesis. In addition, the methyl ester forms of these "alkaloids" were found to be artifacts derived from corresponding acids during soxhlet extraction (Passreiter, 1998). Therefore, the only true alkaloids reported from the Asteraceae are from the Senecioneae, Eupatorieae, and Cynareae.

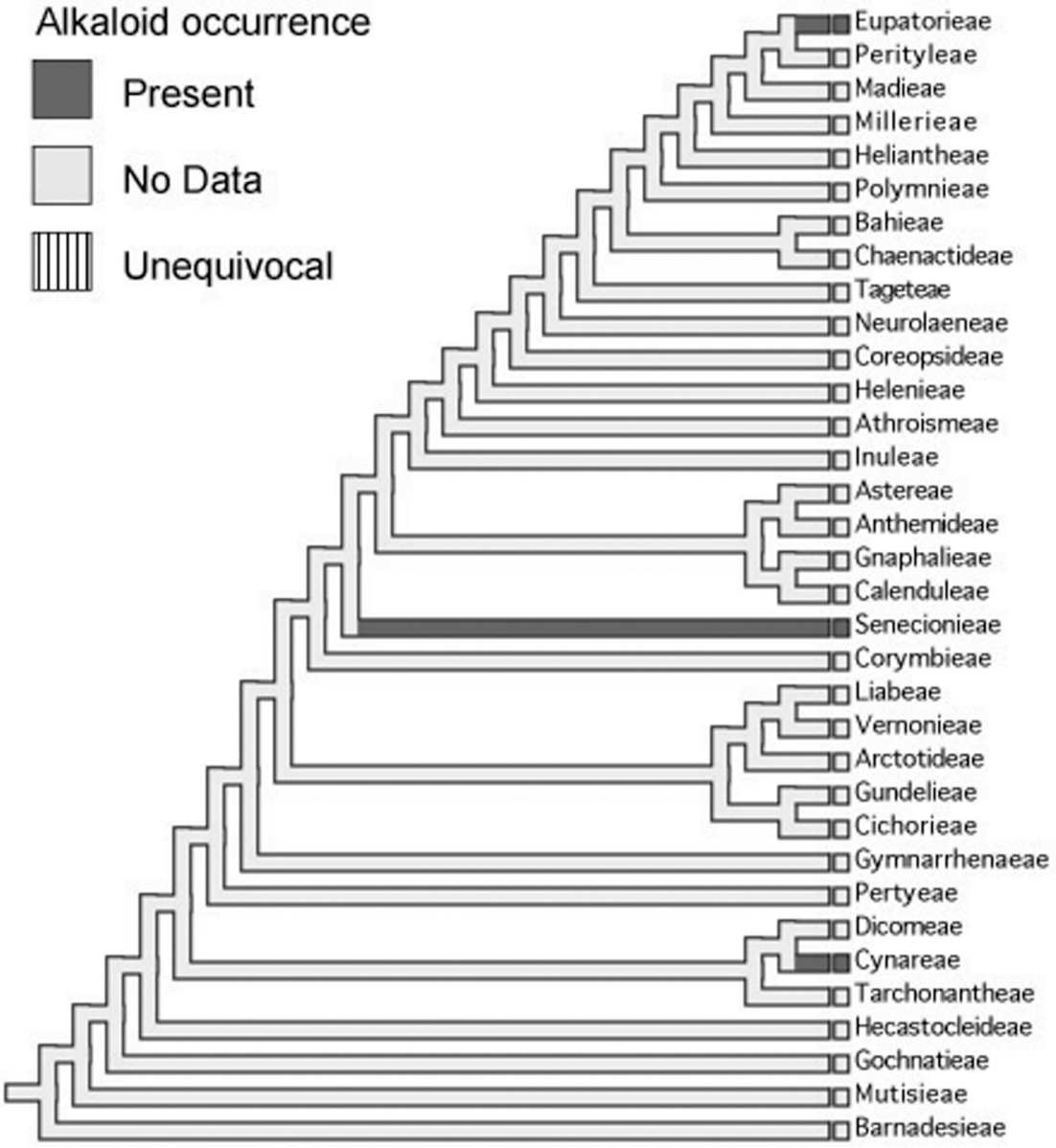
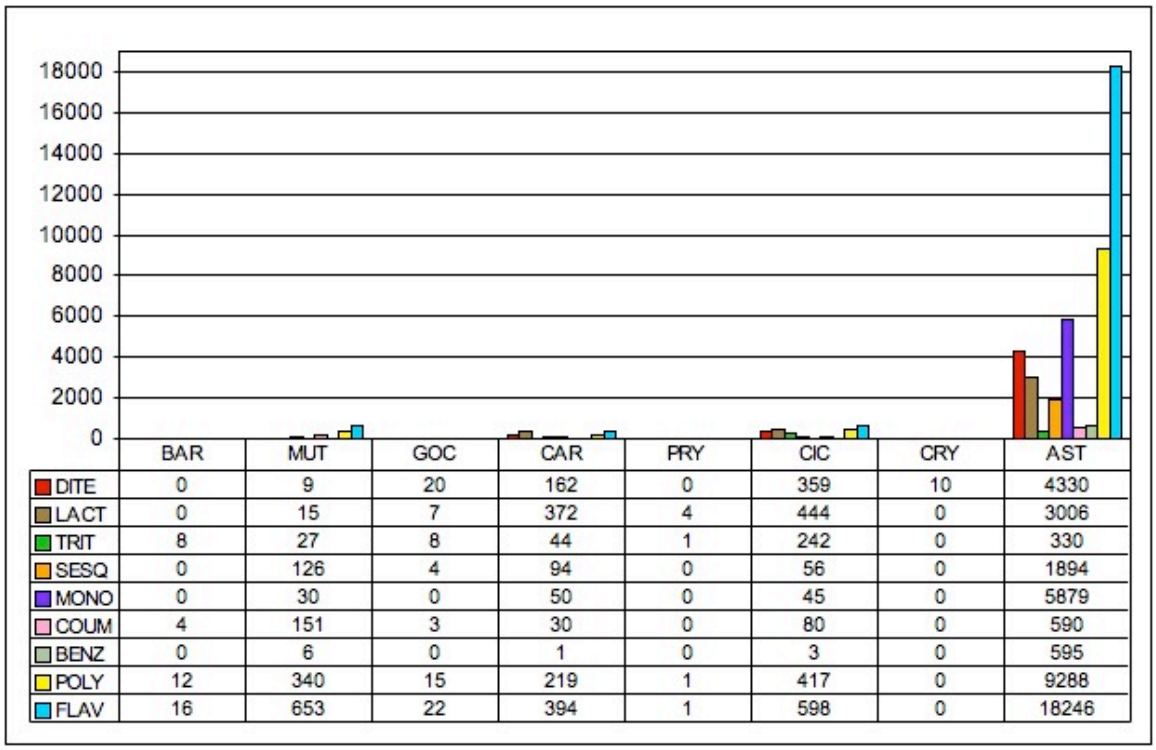


Figure 4.4. Distribution of alkaloids viewed at the tribal level, expressed as "presence" or "absence" of chemical data recorded in our database for the Asteraceae. DNA-based phylogenetic framework adapted from Panero and Funk (2002) and Funk et al. (2005).

The occurrence of pyrrolizidine alkaloids in Senecioneae has been studied in the context of insect co-evolution (Pasteels et al., 2001). The leaf beetles *Platyphora* and *Oreina* spp., (Coleoptera: Chrysomelinae) are known to sequester pyrrolizidine alkaloids from their host plants and to store these compounds as defensive secretions (Pasteels et al., 2001). Interestingly, the authors also reported the presence of saponins in the defensive secretions of *Platyphora*, but saponins were not detected in the leaf beetles food plant, indicating these compounds are synthesized by the beetles, and not sequestered as in the case of the pyrrolizidine alkaloids. Further studies by Plasman et al. (2001) indicated that although the host plants were devoid of saponins, the triterpene precursor β-amyrin, was confirmed to be present in host plant species on which the beetles fed. This suggests that insects of the genus *Platypohora* possess the necessary enzymes for transformation of β-amyrin in their food plant to the oxidized form oleanolic acid and for their subsequent glycosylation (Plasman et al., 2001). Considering that species of the leaf beetle family (with few exceptions) feed exclusively within a single plant family or even a single genus, this example shows the importance of examining chemical profiles for understanding the origin and evolution of plant-insect interactions.

Subfamilies. --Total chemical occurrences for each subfamily and tribe were calculated in Excel and converted into several graphics (Figures 4.5-4.8) showing the distribution and abundance of each chemical class. Figure 4.5 shows the occurrence of secondary metabolites in subfamilies of the Asteraceae according to Panero and Funk (2002) excluding the subfamilies Hecastocleidoideae and Gymnarrhenoideae, for which no chemical data have been reported.



	BAR	MUT	GOC	CAR	PRY	CIC	CRY	AST
DITE	0	9	20	162	0	359	10	4330
LACT	0	15	7	372	4	444	0	3006
TRIT	8	27	8	44	1	242	0	330
SESQ	0	126	4	94	0	56	0	1894
MONO	0	30	0	50	0	45	0	5879
COUM	4	151	3	30	0	80	0	590
BENZ	0	6	0	1	0	3	0	595
POLY	12	340	15	219	1	417	0	9288
FLAV	16	653	22	394	1	598	0	18246

Figure 4.5. Total occurrences for each chemical class found in our chemical database for subfamilies of the Asteraceae according to Panero and Funk (2002).

The monophyletic Barnadesioideae, which are sister to the rest of the Asteraceae, have an extremely simple chemistry characterized by the production of a relatively small number of flavonoids, polyacetylenes, coumarins and triterpenes. Similarly, chemical reviews published for the Calyceraceae and Goodeniaceae, the closest relatives to the Asteraceae, indicate a very simple chemical profile much like that of Barnadesioideae (Bohm et al., 1995; Ghisalberti, 2004). The Barnadesioideae represent only 0.4% (Jeffrey, 2007) of all extant species in the Asteraceae and less than 0.1% of the chemical occurrences in our database. In contrast, the Asteroideae are the largest subfamily in the Asteraceae, containing ~72% (Jeffrey, 2007) of the species in the family and have more chemical occurrences and chemical diversity than all other subfamilies combined (Figure

4.5). Every chemical class is represented in the Asteroideae, with flavonoids reported the most frequently, with over 18,000 chemical occurrences. The monoterpenes are also quite abundant, which coincides with the complex pollination systems and insect interactions in the Asteroideae.

The subfamilies positioned in between the Barnadesioideae and the Asteroideae are clades that historically made up the large subfamily Cichorioideae (28% of the species in the family); more recently the Cichorioideae were divided into several smaller subfamilies (Panero and Funk, 2002). Of these subfamilies, the Mutisioideae, Carduoideae, and Cichorioideae exhibit the most chemical diversity and abundance, as illustrated in Figure 4.5.

Figure 4.6 shows the percentage of each chemical class as individual pie charts for subfamilies Mutisioideae, Carduoideae, Cichorioideae and Asteroideae. Terpenoids (including mono-, di, sesqui-and tri-) account for more than half of the total chemical occurrences in the Cichorioideae (51%) and the Carduoideae (53%), while flavonoids represent 26% of the total chemical occurrences in the Cichorioideae and 29% in the Carduoideae. In contrast, nearly half of the total chemical occurrences for the Asteroideae (42%) and the Mutisioideae (49%) are attributed to flavonoids, while terpenoids represent 15% of the total occurrences in the Mutisioideae and 35% in the Asteroideae.

Polyacetylenes represent the third largest component of the total chemical occurrences in all four major subfamilies (Figure 4.6), ranging from 16% of the total chemical occurrences in the Carduoideae to a maximum of 25% in the Mutisioideae. In the Cichorioideae and in the Asteroideae polyacetylenes represent 19% and 21% of the

total chemical occurrences. Benzofurans account for less than 1% of total chemical occurrences in all subfamilies, except for the Asteroideae, where total occurrences slightly exceeds 1%. The Mutisioideae contain an unusually large number of coumarins in comparison with other subfamilies of the Asteraceae, representing 11% of the total chemical occurrences, whereas coumarins represent 4% of the total occurrences in the Cichorioideae, 2% in the Cardueae and only 1% in the Asteroideae.

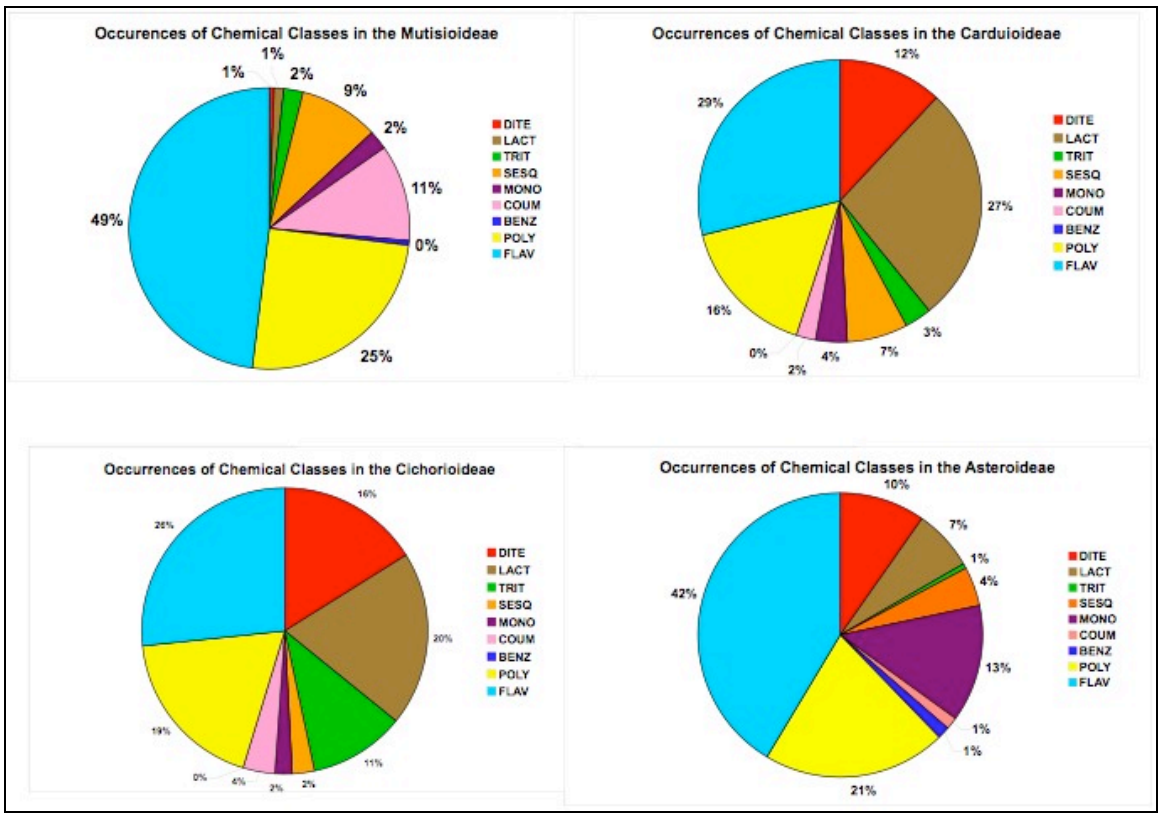


Figure 4.6. Individual pie charts for subfamilies Mutisioideae, Carduoideae, Cichorioideae and Asteroideae representing the total chemical occurrences expressed as percentages of each chemical class found in our chemical database for the Asteraceae.

The chemical data presented in Figure 4.6 illustrate how easily plants can adjust their production of secondary metabolites to adapt to changing environments. Each subfamily exhibits a distinct chemical ratio, characterized by both the expansion and reduction of different chemical classes. The ancestral state of secondary metabolite expression in the Asteraceae is not known and cannot be inferred from extant species, however, given the phylogenetic relationships, one can infer a direction in which the chemical profiles are changing. Considering that terpenoids are more expensive to manufacture per gram than most other primary and secondary metabolites (Gershenzon, 1994), the benefits of devoting more than half of the total biosynthetic expenditures for secondary metabolism to produce a single class of compounds must outweigh the cost of manufacturing this one class of compounds for taxa in the Cichorioideae and Carduoideae. On the other hand, the same must be true for the production of flavonoids in the Mutisioideae and Asteroideae.

The observed shifts in metabolism between subfamilies are most likely the result of natural selection. However, a single genetic mutation can sometimes drastically change the chemical properties of a plant, (i.e. taste, toxicity, etc). The diversity and abundance of a particular chemical class could simply be artifacts of mutated enzymes with promiscuous behavior that may or may not benefit plant fitness. Furthermore, because our data are compiled from literature reports, many of which represent studies on bioactive and medicinally important compounds, we cannot be certain that the chemical occurrence patterns described here represent their natural distribution in the plants. Still, further attempts to explain these patterns should stimulate many interesting research

questions that could be addressed using a combination of phytochemical, genetic and ecological data.

Tribes. --The distribution and abundance of each chemical class can also be viewed at the tribal level, as illustrated in Figures 4.7 and 4.8. The maximum abundance of chemical classes is seen in the Anthemideae, due to the accumulation of large amounts of flavonoids, monoterpenes and polyacetylenes and to the intensity of research devoted to medicinally important taxa from this group such as *Chrysanthemum, Achillea, Artemisia* and *Anthemis* (Teixeira da Silva et al, 2004, Teixeira da Silva, 2004). When the Anthemideae is excluded, the tribe with the most chemical occurrences is the Astereae, as illustrated in Figure 4.8.

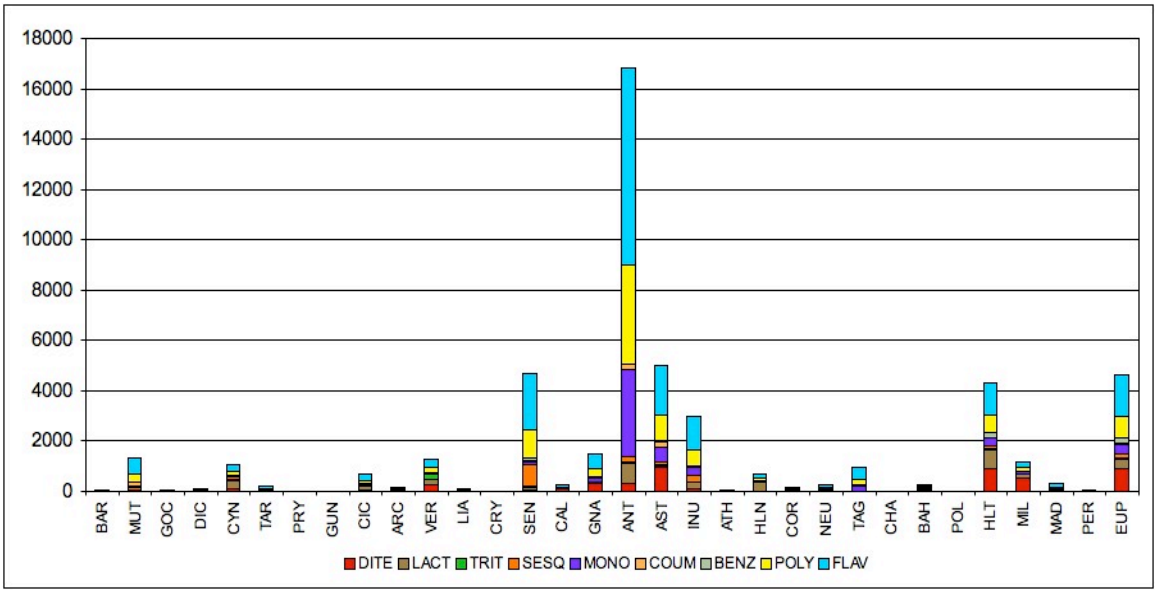


Figure 4.7. Total occurrences of each chemical class found in our chemical database for tribes of the Asteraceae. Genera were assigned to tribes according to Panero and Funk (2002), Funk et al. (2005) and Panero (2007).

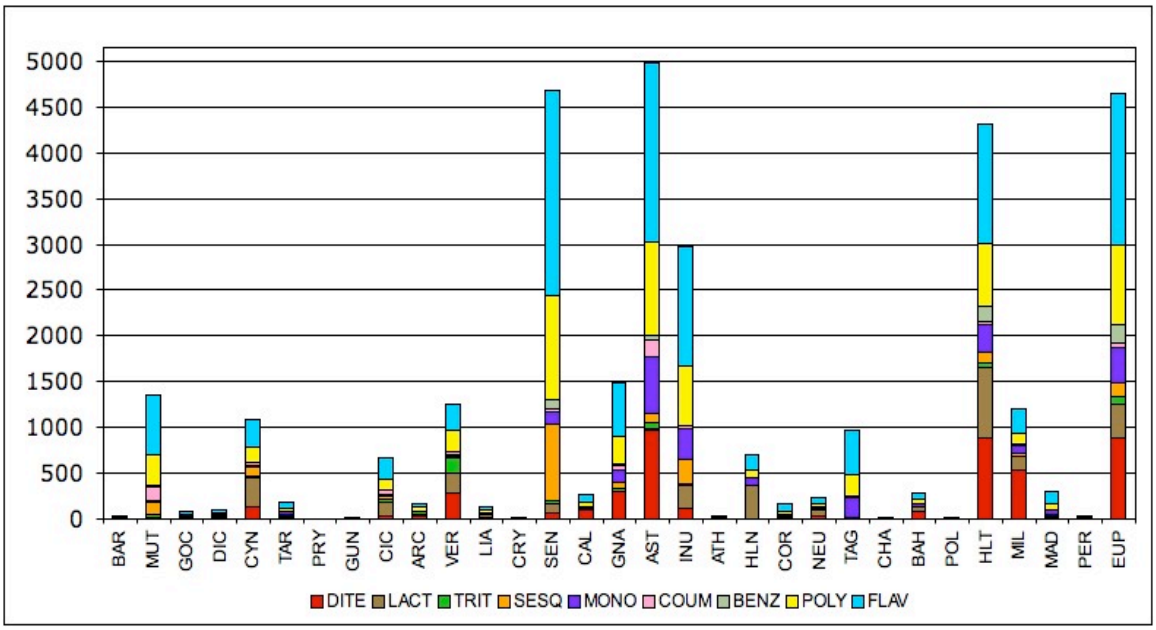


Figure 4.8. Total occurrences of each chemical class found in our chemical database for tribes of the Asteraceae, excluding Anthemideae. Genera were assigned to tribes according to Panero & Funk (2002), Funk and al. (2005) and Panero (2007).

Flavonoids account for 2,000 of the ~5,000 chemical occurrences in the Astereae. Polyacetylenes and diterpenes are also quite abundant in the Astereae, with ~1000 chemical occurrences. All other chemical classes are represented in the Astereae but in much smaller quantities then the above-mentioned classes (Figure 4.8). The Senecioneae and Eupatorieae both accumulate large amounts of flavonoids and diterpenes, but differ in their expression of sesquiterpenes and diterpenes. While the Senecioneae tends to accumulate sesquiterpenes, the Eupatorieae instead accumulate diterpenes.

The Heliantheae show slightly fewer chemical occurrences than the Eupatorieae and Senecioneae, in ratios similar to the Eupatorieae. Finally, the Inuleae also accumulate large amounts of secondary metabolites, mainly flavonoids and polyacetylenes and to a lesser extent, the mono-, di, sesqui- and sesquiterpene lactones.

All other tribes have fewer then 1500 chemical occurrences reported. The Mutisieae are of particular interest because of their basal position in the family, their problematic classification and diverse chemical profiles characterized by complex coumarins and accumulation of flavonoids, polyacetylenes, and sesquiterpenes. The Vernonieae also deserve mention because they accumulate a large number of diverse triterpene skeletons.

Subtribe and genera level views of triterpenes saponins in the Asteraceae.-- The genus *Silphium* L. (tribe Heliantheae) consists of 11 species native to North America, distributed primarily in the eastern United States and extending into southeastern Canada. Recently, the phylogenetic relationship of *Silphium* and subtribe Engelmanniinae were examined using DNA sequence data (Clevinger and Panero, 2000; Panero, 2007). The resulting phylogeny (Figures 4.9) will be employed here to summarize the distribution of triterpene saponins at the subtribal, generic, and species level. The oxidation patterns of triterpenes from the genus *Silphium* will also be discussed.

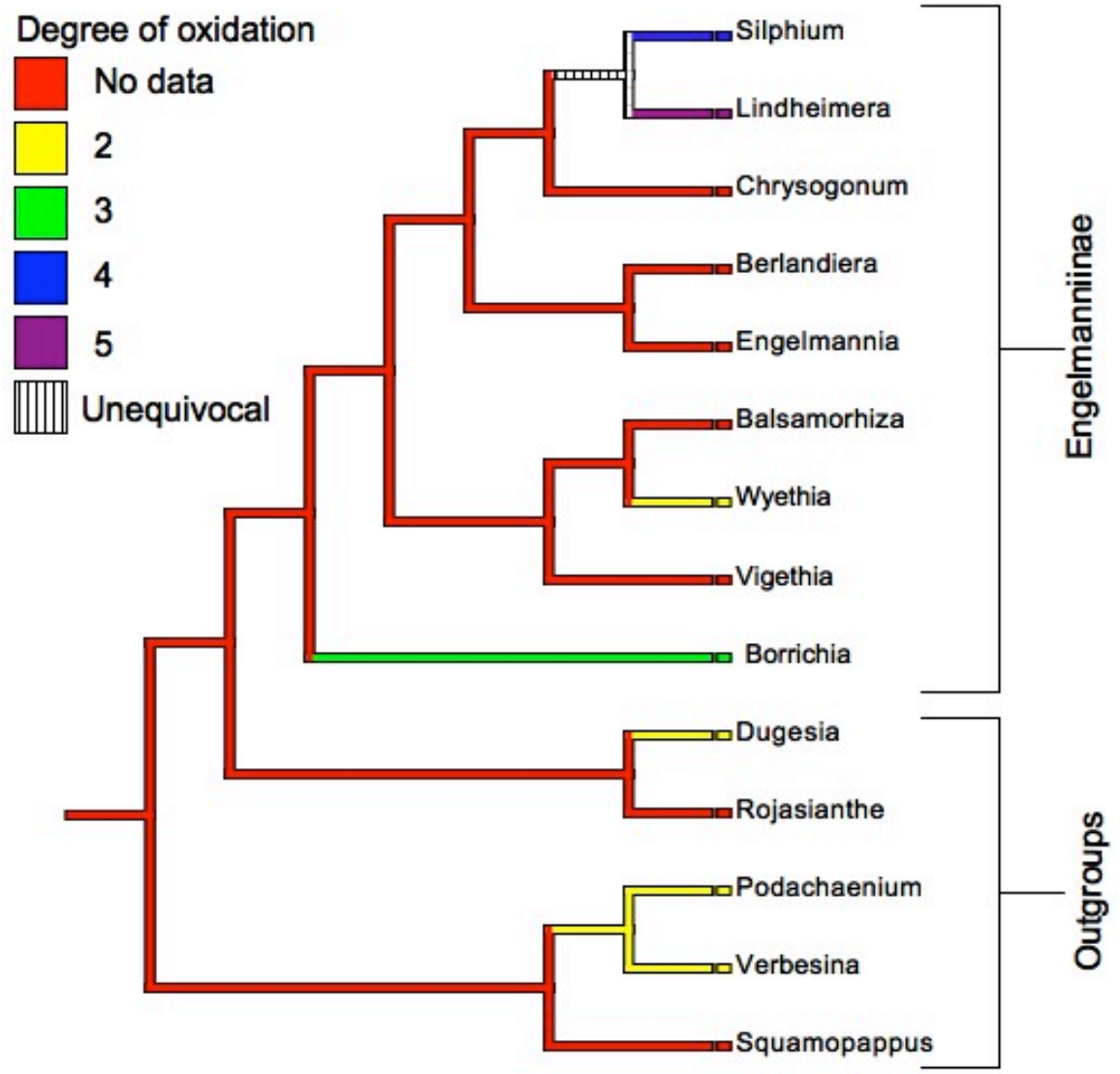


Figure 4.9. Degree of triterpene oxidation present in the subtribe Engelmanniinae and closely related genera. DNA-based phylogenetic framework adapted from Clevinger and Panero (2000) with subtribal delineations updated by Panero (2007).

Silphium is the most extensively studied taxon in the subtribe Engelmanniinae, due in part to its common use as a folk medicine by several Native American tribes. In recent years, our research group at UT-Austin has focused primarily on documenting the

flavonoids, phenolic acids and, more recently, the triterpene-type saponins in all species of *Silphium* (El-Sayed et al., 2002; Wojcinska et al., 2006).

Until recently, the utilization of triterpenes and their glycosidic forms, the saponins, in chemosystematic studies have been limited. One recently published excellent review outlines the evolutionary development and distribution of these compounds in higher orders of angiosperms and summarizes findings on structural and distributional data obtained for triterpenoid saponins during the last 50 years (Henry, 2005). Based on these data, it is clear that saponins are restricted to advanced taxonomic groups such as the Carophyllideae, primitive Rosideae and Asterideae, but are lacking in the Paleodicots and Monocots. The only exception thus far was reported by Osbourn's group (Papadopoulou et al., 1999), who used a combination of phytochemical, molecular and genetic techniques to examine the expression patterns of naturally-occurring triterpene saponins found in oats (*Avena* spp.); these compounds proved critical in defense against pathogenic fungi.

As suggested by Gottlieb and coworkers (1989) in their evolutionary studies of secondary metabolism in Angiosperms, the degree of oxidation can be correlated with evolutionary advancement, with highly oxidized compounds indicating a greater capacity to utilize atmospheric oxygen (see section 4.2). Therefore, you would expect to see more highly oxidized compounds in more derived lineages. Early biosynthetic intermediates in the triterpene pathway, such as oleanolic acid, are represented in much greater numbers throughout the angiosperm lineages than later metabolic intermediates such as quilliac acid implying a direction of evolution of chemical pathways in angiosperm lineages, with

later biosynthetic intermediates more oxidized and primarily restricted to the more advanced groups of angiosperms, such as the Asteraceae (Henry, 2005). Because researchers now understand some of the main biosynthetic routes and intermediates in the production of saponins, quantitative and qualitative analyses of these different intermediates could help to clarify the role of saponins in the evolution of Asteraceae family.

Through our investigation of triterpene saponins from *Silphium* we have found a range of structural types and patterns of occurrences. Some species display a rich mixture of saponins, while others contain only a few, structurally simple compounds. However, it is clear from our preliminary data that all genera of the subtribe Engelmanniinae produce saponins (data unpublished). Our results from a detailed structural analysis of triterpenes from *Silphium radula* Nutt., which belongs to the most derived clade in the genus *Silphium*, support this hypothesis based on the content of several novel triterpene saponins with high degrees of oxidation, including hydroxyl, carboxyl and aldehyde functional groups (Calabria et al., 2008). Figure 4.9 illustrates this concept showing our preliminary data on the distribution and occurrence of oxidized triterpenes in the subtribe Engelmanniinae. Although the occurrence of triterpenes in all genera in the subtribe was confirmed through our chemical investigations, a complete data set including oxidation states is still underway. Therefore, Figure 4.9 shows our current knowledge of the degree of oxidation present in the subtribe Engelmanniinae and the closely related genera, *Dugesia, Rojasianthe*, *Podachenium*, *Verbesina* and *Squamopappus.* The highest degree of triterpene oxidation is seen in *Lindheimera* and *Silphium,* where 4 to 5 carbons are

oxidized. *Borrichia* shows 3 carbons oxidized and *Wyethia, Dugesia, Podachenium* and *Verbesina* show the least degree of oxidation with only 2 groups oxidized. These preliminary data support the hypothesis that highly oxidized triterpenes occur in more derived taxa at the subtribal level. Further investigations of saponins at the generic and species levels are currently underway.

Species level.-- Several species of gall wasps (Hymenoptera: Cynipidae) induce galls on *Silphium*. Gall-inducing insects are typically specialist plant feeders with most species confined to one specific host plant species. An excellent example involves several gall-inducing wasp species of the genus *Antistrophus* whose larvae induce galls in the pith and cambium of stems of *Silphium* spp. (Tooker and Hanks, 2006). Of the ten or so species of *Antistrophus* native to North America, at least five form galls in either the stems or flowering heads of *Silphium* species (Tooker and Hanks, 2004). For stem-inhabiting species, female gall wasps lay eggs in bolting stems and feeding by newly hatched larvae induces gall formation. A specific blend of monoterpenes emitted from the host plant is critical for female wasps to recognize their specific host plants (Tooker et al., 2005). In addition, upon emergence from galls, adult males use olfactory signals from their natal host plant species to locate their mates nested within the dead over-wintering stems of *Silphium.* This same blend of volatiles emitted by *Silphium* also attracts the parasitic wasp *Eurytroma lutea* Bugbee (Hymenoptera: Eurytomidae), which is a natural enemy of *Antistrophus*. The larvae of *E. lutea* kill gall wasp larvae, diminishing the amount of plant damage inflicted by *Antistrophus* and allowing the plants to put resources into reproduction that would have otherwise been consumed by the gall wasp

larvae. As a result, galled *S. laciniatum* plants that are able to attract *E. lutea* appear to have improved fitness relative to galled controlled plants (Tooker and Hanks, 2006).

The molecular and biochemical bases for these interactions are largely unstudied. However, the manipulation of host plants chemistry by galling (and many other) insects is well documented. Modes of action range from the induction of the hypersensitive response, suppression of genes involved in biosynthesis of defensive compounds, detoxification of active defense compounds and even tissue-specific redistribution of secondary metabolites, to name a few. For example, the nutritive tissues lining galls often are void of secondary metabolites, whereas the exterior of the galls can have relatively high concentrations of toxins. This distribution of secondary compounds gives galling insects a toxin-free food source while providing the gall insects protection against natural enemies or herbivores that might favor plant tissue (personal communication, Tooker).

The story of *Silphium* and the gall wasp *Antistrophus* demonstrates how species level chemical data can be useful for understanding plant-insect interactions in the Asteraceae. To date, *Antistrophus* galls have been documented in four of the eleven species in the genus *Silphium*; the question of why these four species are the preferred host for *Antistrophus* remains to be answered. The complex tritrophic interactions described here are mediated by chemical signals produced by *Silphium* and manipulated by insects in the surrounding environment. It is our own view that research utilizing both phytochemical and molecular systematic tools may help to explain such interactions. In addition, mapping chemical characters on host-plant molecular phylogenies coupled with

phylogenies of their associated insects could also help to elucidate macroevolutionary patterns in both plant and insect lineages.

4.5 Conclusions

In summary, we have presented a historical overview of studies on the secondary chemistry of the Asteraceae and outlined the current distribution, abundance, and diversity of secondary metabolites in the context of DNA-based phylogenetic trees for the family (Clevinger and Panero, 2000; Panero and Funk, 2002; Funk et al., 2005; Panero, 2007). The examples described here demonstrate the usefulness of chemical data for addressing evolutionary questions at different levels of order, both in terms of secondary metabolite data and different taxonomic levels.

The chemical data summarized here indicate that although there are thousands of secondary metabolites described from nearly every tribe of the Asteraceae, there are still many questions that could be addressed if additional chemical data were available. Furthermore, model species from the Asteraceae with diverse chemical profiles are desperately needed to correlate phytochemical findings with genetic and genomic data for secondary metabolic pathways. This vision is slowly becoming a reality through the recent efforts of the Compositae Genome Project (Rieseburg and Michelmore, 2003), which provides genomic resources and tools for the Asteraceae, and other researcher groups developing model Asteraceae such as the Gerbera laboratory (Teeri et al., 2006). Considering that ~170 secondary metabolites from 7 major chemical classes have been identified from the model plant *Arabidopsis thaliana* (L.) Heynh, representing a five-fold

increase over the last 10 years (D'Auria and Gershenzon, 2005), one can imagine the potential for discovering even more secondary metabolite diversity in the Asteraceae as more genomic information becomes available. With these resources, molecular biologists, chemists, ecologists, and systematists should work together to address questions about the functional importance of secondary metabolites in the Asteraceae in the context of their life histories. And although DNA data provides the most reliable method for estimating evolutionary relationships and distances between taxa, these data cannot explain how or why a particular plant evolved without phenotypic information, including a broad range of morphological and chemical characters. This is an exciting time to be involved with investigations of the secondary chemistry of the Asteraceae, and future studies will almost certainly shed light on the previously unknown roles of secondary metabolites in the evolution and diversification of this family.

Chapter 5: Isolation and Characterization of Triterpene Saponins from *Silphium*

5.1 Triterpene saponins from *Silphium radula*[3]

Nine new triterpene saponins (**1-9**) were isolated from the leaves and stems of *Silphium radula* Nutt. (Asteraceae). Their structures were determined by extensive 1D (^{13}C, ^{1}H, DEPT, TOCSY) and 2D NMR (NOESY, HSQC, HMBC) and ESI-MS studies. The compounds were identified as 3β,6β,16β-trihydroxyolean-12-en-23-al-3-*O*-β-glucopyranosyl-16-*O*-β-glucopyranoside (**1**), urs-12-ene-3β,6β,16β-triol-3-*O*-β-galactopyranosyl-(1 → 2)-β-glucopyranoside (**2**), 3β,6β,16β-trihydroxyolean-12-en-23-oic acid-3-*O*-β-glucopyranosyl-16-*O*-β-glucopyranoside (**3**), urs-12-ene-3β,6β,16β,21β-tetraol-3-*O*-β-glucopyranoside (**4**), olean-12-ene-3β,6β,16β,21β-tetraol-3-*O*-β-glucopyranoside (**5**), olean-12-ene-3β,6β,16β,21β,23-pentaol-3-*O*-β-glucopyranosyl-16-*O*-β-glucopyranoside (**6**), olean-12-ene-3β,6β,16β-triol-3-*O*-β-glucopyranosyl-16-*O*-α-arabinopyranosyl-(1 → 2)-β-glucopyranoside (**7**), olean-12-ene-3β,6β,16β,23-tetraol-3-*O*-β-glucopyranosyl-16-*O*-α-arabinopyranosyl-(1 → 2)-β-glucopyranoside (**8**), 3β,6β,16β,21β-tetrahydroxyolean-12-en-23-al-3-*O*-β-glucopyranoside (**9**). The presence of a 6β–hydroxyl function was not common in the oleanene or ursene class and the aglycones of these compounds were not found previously in the literature. Moreover, the cytotoxic activities of the isolated compounds were tested against human breast cancer

[3] This chapter was originally published as: "Triterpene saponins from *Silphium radula"*
Lalita M. Calabria, Sonia Piacente, Ireneusz Kapusta, Suranganie F. Dharmawardhane, Frances M. Segarra, Peter J. Pessiki, Tom J. Mabry. 2008. Phytochemistry 69, 961-972.

cell line MDA-MB-231. Results showed that compound **2** decreased cell proliferation in a statistically significant manner at 25 µg/ml.

5.1.1 Introduction

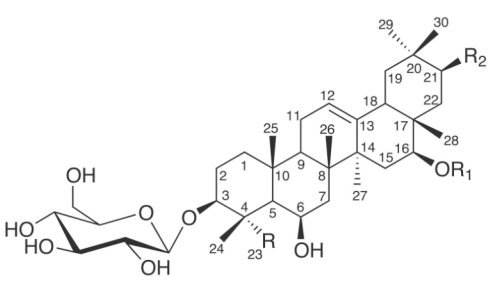
Silphium L. (Asteraceae) is a small genus of 11 species common to the prairies and woodlands of central North America (Clevinger and Panero, 2000). Native American tribes, including the Cherokee Indians, considered *Silphium* a medicinal herb and prepared extracts of the leaves and flowers for respiratory and kidney ailments and also applied the roots of *Silphium* as a poultice for bleeding wounds, backaches and hemorrhaging (Hamel and Chiltoskey, 1975). The potential of *Silphium* as a natural remedy is also well known in parts of Europe, where it is cultivated as a garden ornamental and herbal medicine. Several chemical investigations of European-grown *Silphium* species including *S. perfoliatum, S. integrifolium* and *S. trifoliatum* (now considered a *subspecies of S. asteriscus*, Clevinger and Panero, 2000) have appeared in the literature in recent years (Kowalski, 2004). For example, chemical analysis of *Silphium perfoliatium* resulted in the isolation of flavonoids (El-Sayed et al., 2002), phenolic acids (Kowalski and Wolski, 2003), essential oils (Kowalski et al., 2005), sesquiterpenes (Bohlmann and Jakupovic, 1979; 1980), diterpenes (Pcolinski et al., 1994), and oleanolic-type triterpene saponins (Davidyants et al., 1984a,b,c; Davidyants et al., 1985; Davidyants et al., 1986). The saponin mixture showed blood cholesterol-lowering activity as well as fungicidal properties (Syrov et al., 1992; Davidyants et al., 1997). In addition, there has been an increasing interest in European-cultivated species of

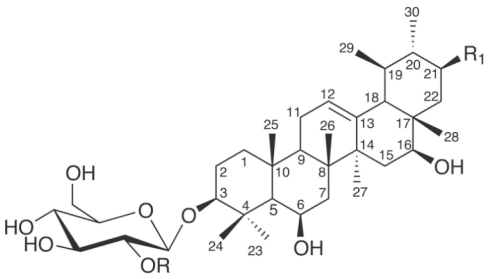
Silphium as commercial sources for oleanolic and ursolic-type triterpenes (Kowalski, 2007).

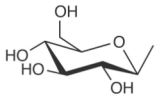
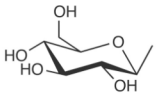
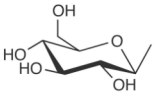
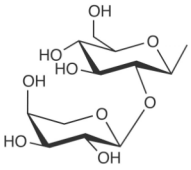
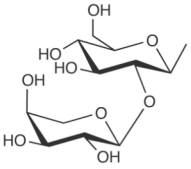
In an effort to further characterize triterpene saponins of chemosystematic and medicinal significance from *Silphium*, our research group has undertaken chemical investigations of all *Silphium* species harvested from their native habitats across North America. We report here the isolation and structural elucidation of nine new pentacyclic oleanene and ursene-type triterpene saponins (**1-9**) from leaf and stem methanolic extracts of *Silphium radula* Nutt., a species found in prairie remnants and rocky open forests of the southeastern United States (Rickett, 1967). To our knowledge, this is the first chemical report for this species.

5.1.2 Results and Discussion

The concentrated methanolic extract of a mixture of *S. radula* leaves and stems was subjected to solid phase extraction on reverse phase C-18, CC on silica gel and MPLC on reversed phase C-18, affording nine new saponins (**1-9**), shown in Figure 5.1. Their structures were established by 1D and 2D NMR spectroscopy and ESI-MS analyses.





	R	R_1	R_2
1	CHO		H
3	COOH		H
5	CH_3	H	OH
6	CH_2OH		OH
7	CH_3		H
8	CH_2OH		H
9	CHO	H	OH

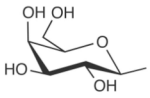
	R	R_1
2		H
4	H	OH

Figure 5.1. Structures of polyhydroxylated triterpene saponins (1-9) isolated from *Silphium radula*.

The HR-ESI-MS (negative-ion mode) of compound **1** exhibited a pseudomolecular ion peak at *m/z* 795.4531 $[M–H]^-$, ascribable to a molecular formula of $C_{42}H_{68}O_{14}$. The ESI-MS experiment gave (negative-ion mode) a quasi-molecular ion peak at *m/z* 795 $[M–H]^-$ indicating a molecular weight of 796 for compound **1**. Further fragment ion peaks in the ESI-MS-MS spectrum were observed at *m/z* 633 $[M–H\text{-}162]^-$ and *m/z* 472 $[M–H\text{-}162\text{-}162]^-$corresponding to the successive loss of two hexosyl moieties. This result suggested that saponin **1** contained two sugar units. The ^{13}C-NMR spectrum of compound **1** showed 42 signals, of which 30 were assigned to a triterpenoid moiety and 12 to the saccharide portion. The ^{1}H-NMR spectrum of the aglycone moiety of **1** showed signals for seven tertiary methyl groups (δ 0.95, 0.98, 1.15, 1.44, 1.51, 1.56, 2.00), along with the signals at δ 5.38 (1H, t, J = 3.7 Hz) and 9.87 (1H, s) ascribable, respectively, to an olefinic and an aldehydic proton (Figure 5.2). Further features were signals at δ 4.13 (1H, br s), 4.22 (1H, dd, J = 11.0, 4.5 Hz), and 4.58 (1H, dd, J = 11.5, 4.5 Hz), indicative of secondary alcoholic functions (Table A.1, see Appendix A).

In the ^{13}C-NMR spectrum the signals at δ 122.6 and 143.5 ascribable to C-12 and C-13 suggested a Δ^{12} oleanene skeleton (Mahato and Kundu, 1994). Full assignments of the proton and carbon resonances of the aglycone (Table A.1; Appendix A) were secured by ^{1}H-^{1}H DQF-COSY and HSQC spectra. The aldehyde function was located at C-23 on the basis of the downfield shift exhibited by C-4 (δ 56.4) and the highfield shifts exhibited, respectively, by C-3 (δ 82.3), C-5 (δ 49.3), and C-24 (δ 11.9) in comparison with the same carbon resonances in an oleanene skeleton bearing a Me-23 (Mahato and Kundu, 1994).

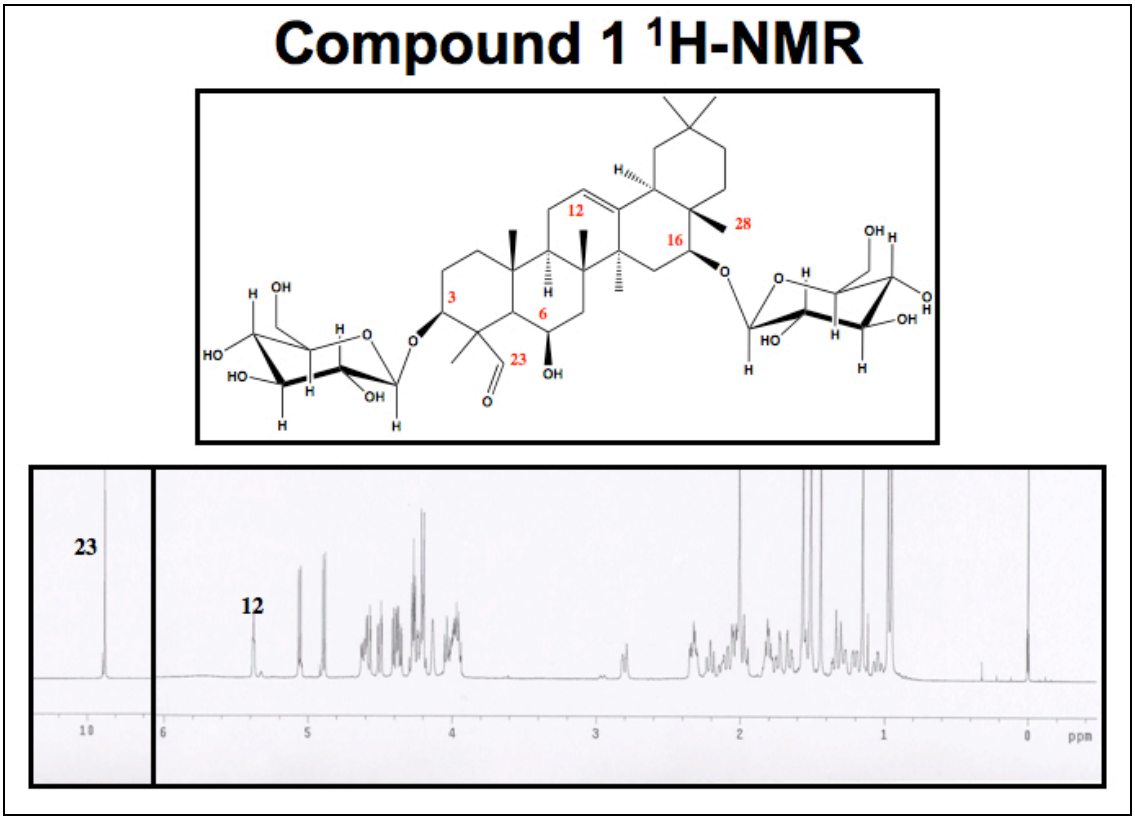


Figure 5.2. H-NMR spectra and structure for compound 1 isolated from *Silphium radula*, highlighting the aldehydic and olefinic proton resonances.

Also, the chemical shift of Me-24 in the ^{1}H-NMR spectrum (δ_H 2.00) was diagnostic for a 23-CHO (De Tommasi et al., 1998). The HMBC spectrum of **1** confirmed the position of the aldehyde function showing significant cross-peaks, due to $^{2}J_{C\text{-}H}$ and $^{3}J_{C\text{-}H}$ correlations between H-23 (δ 9.87) and C-3 (δ 82.3), C-4 (δ 56.4), and C-24 (δ 11.9). The signal at δ 4.22 (dd, J = 11.0, 4.5 Hz) attributed to H-3ax suggested the presence of a β-OH group at C-3. The location of a hydroxyl group at C-6 was determined by the correlation between H-6 (δ 4.13, br s) and C-5 (δ 56.4). H-5 α and H-6 resonated as broad singlets at δ 1.50 and 4.13, respectively, indicating their axial-equatorial relationship. Thus, the β-orientation of the C-6 hydroxyl group was defined (Wang et al., 2006). Furthermore, a

NOESY correlation between the signal at δ 4.13 (H-6) and the signals 4.22 (H-3) and 9.87 (Me-23) confirmed the α-orientation of H-3 and H-6. The third secondary alcoholic function was located at C-16 on the basis of the carbon resonances of ring D and of the HMBC correlation between the signal at δ 1.15 (Me-28) and the carbon resonance at δ 76.4 (C-16). The coupling constants of H-16 (δ 4.58, dd, J = 11.5, 4.5 Hz) suggested the β-orientation of the C-16 hydroxyl group (Ukiya et al., 2001), which is in good agreement with the nOe effect between H-16 (δ 4.58) and Me-27 (δ 1.44) observed in the NOESY spectrum.

Thus, the aglycone of **1** was identified as 3β,6β,16β-trihydroxyolean-12-en-23-al. Glycosidation of the alcoholic function at C-3 and C-16 were indicated by the downfield shift (~10 ppm) observed for these carbon resonances in **1**, if compared to the corresponding signals in unglycosidated model compounds (Mahato and Kundu, 1994).

For the sugar portion, the ^{1}H-NMR spectrum of **1** (Figure 5.3), showed two anomeric proton signals (δ 4.89, d, J = 7.5 Hz; δ 5.06, d, J = 7.5 Hz). 1D-TOCSY spectra obtained by selectively irradiating anomeric signals gave in both cases the spin-system of a β-glucopyranosyl.

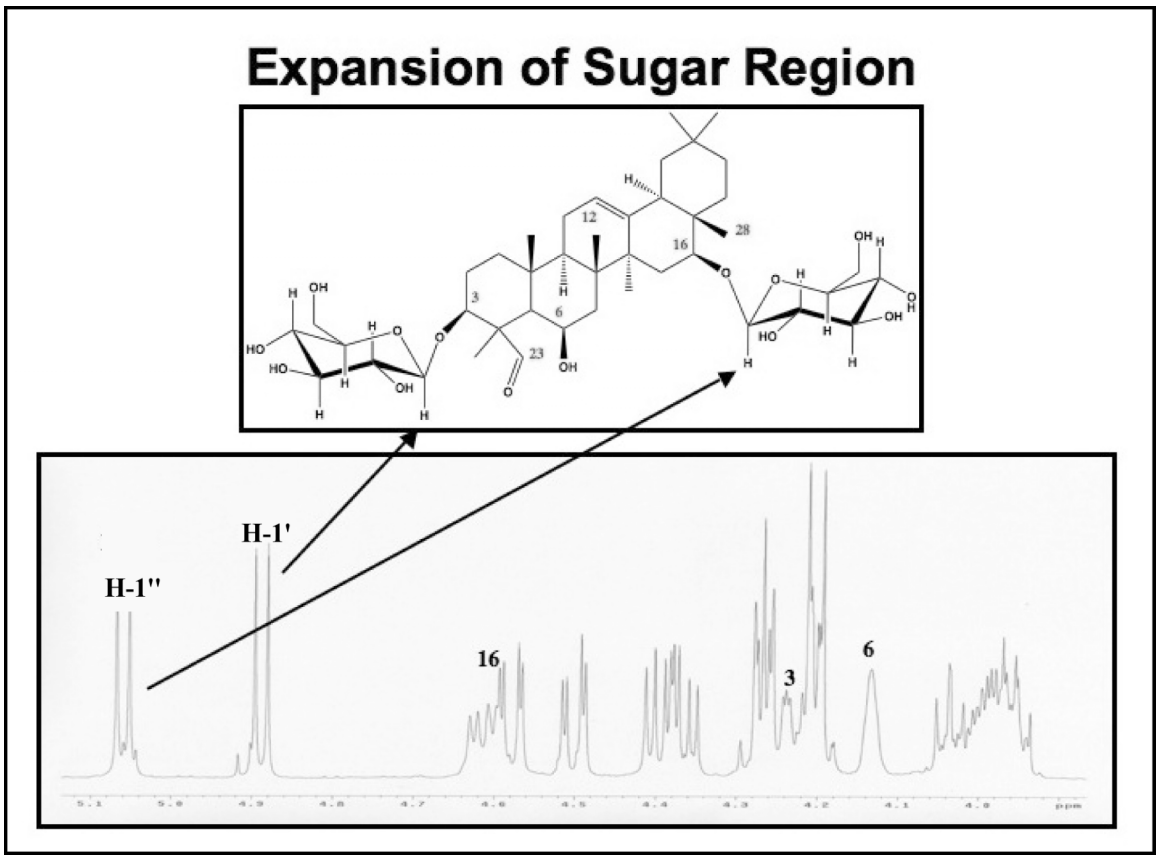


Figure 5.3. H-NMR spectra for compound 1 isolated from *Silphium radula*, showing the expansion of the sugar region. Anomeric protons for the C-16 and C-3 linked glucose residues are shown, in addition to the glycosidated downfield shifts of H-16, H-3 and the hydroxylated broad singlet of H-6.

HSQC experiments, which correlated all the proton resonances with those of each corresponding carbon, showed no glycosidation shifts, suggesting that the two sugars were terminal units (Figure 5.4). The location of the sugar residues at C-3 and C-16 were unambiguously defined by the HMBC experiment, which showed long-range correlations between C-3 (δ 82.3) and H-1glcI (δ 4.89) and C-16 (δ 76.4) and H-1glcII (δ 5.06) (Table A.1; Appendix A). On the basis of the above data compound **1** was identified as 3β,6β,16β-trihydroxyolean-12-en-23-al-3-*O*-β-glucopyranosyl-16-*O*-β-glucopyranoside.

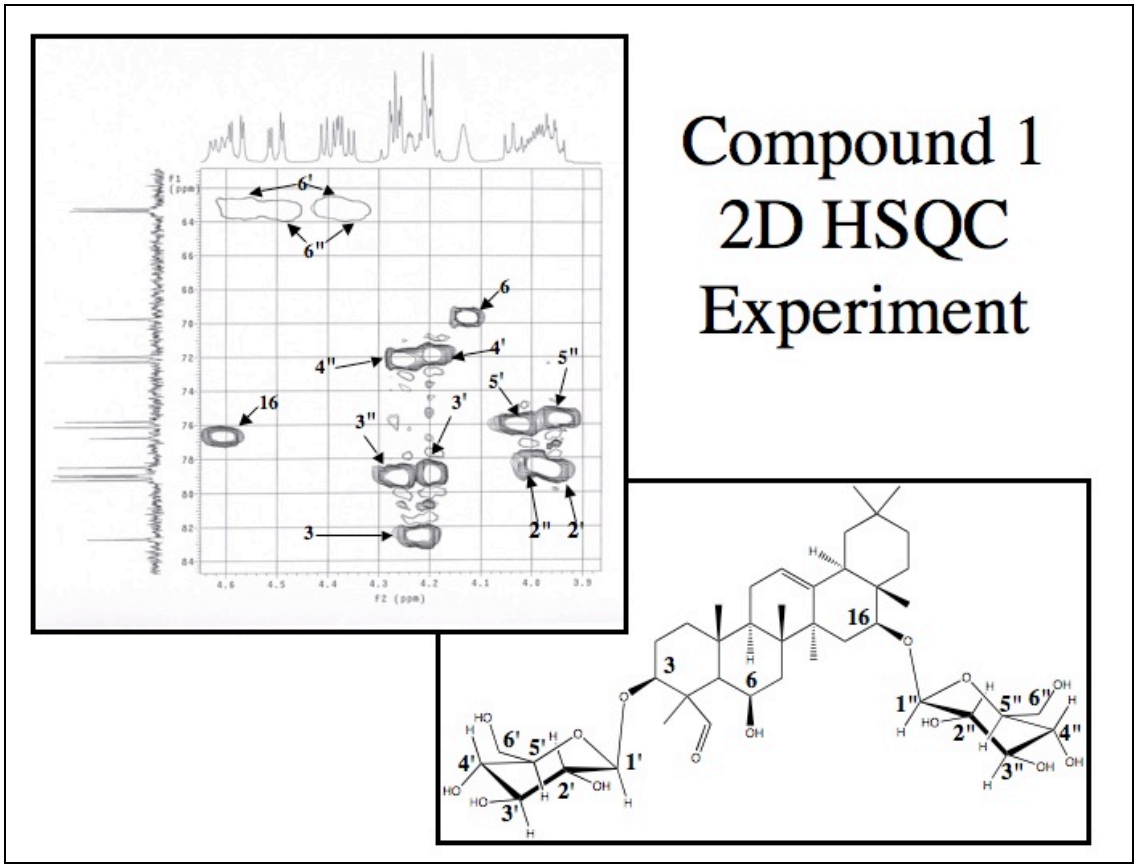


Figure 5.4. HSQC experiment for compound 1 isolated from *Silphium radula*. The carbon-hydrogen correlations for all sugar resonances, glycosidation sites and free hydroxyls are shown.

The negative HR-ESI-MS of compound **2** showed a molecular ion peak at *m/z* 781.4752 $[M–H]^-$, in accordance with an empirical molecular formula of $C_{32}H_{70}O_{13}$. The negative ESI-MS spectrum showed a psuedomolecular ion peak at 781 $[M–H]^-$ indicating a molecular weight of 782 for compound **2**. ESI-MS-MS experiments showed further fragment ions at 619 $[M–H–162]^-$ and 457$[M–H–162–162]^-$, suggesting the presence of two hexosyl moieties. Compound **2** displayed 42 carbon resonances in its ^{13}C NMR spectrum (Table A.1; Appendix A), of which 30 could be assigned to the aglycone and 12

to the sugar moiety. The ^{13}C NMR spectrum of the aglycone portion exhibited eight methyl resonances at δ 17.4, 18.0, 18.6, 18.8, 21.7, 23.0, 25.0, 28.1, two sp^2-hybridized carbons at δ 125.5 and 138.5, resonances for three hydroxymethine groups (δ 65.7, 67.1 and 89.4). These data, when correlated with information from the ^{1}H NMR [six methyl singlets at δ 1.33, 1.45, 1.54, 1.62, 1.66, 1.84, two methyl doublets at δ 0.98 (J = 6.5 Hz) and 0.99 (J = 6.5 Hz) one olefinic proton at δ 5.38, and three oxymethyne protons at δ 3.38, 4.55 and 4.79], indicated that the aglycone of **2** is based on a olean-12-ene-3,6,16-tetraol skeleton. The orientations of the hydroxy groups were assigned as 3β,6β,16β on the basis of ^{1}H NMR coupling constants and by comparison with compound **1**. Assignments of all chemical shifts of protons and carbons of the aglycone portion were ascertained from a combination of 1D-TOCSY, DQF-COSY, HSQC and HMBC analysis. The ^{1}H NMR spectrum for the sugar portion of compound **2** showed two anomeric proton signals at δ 4.98 (1H, d, J = 7.5 Hz) and 5.25 (1H, d, J = 7.5 Hz). The chemical shifts of all the individual protons of the two sugar units were attributed on the basis of 2D-TOCSY and DQF-COSY spectral analysis, and the ^{13}C chemical shifts of their relative attached carbons were clearly assigned from the HSQC spectrum. These data showed the presence of a terminal β-galactopyranosyl (δ 5.25) and a 2-substituted β-glucopyranosyl (δ 4.98) as indicated by the downfield shift of C-2_{glc} (δ 84.6) signal. An unambiguous determination of the linkage site was obtained from the HMBC spectrum, which showed key correlation peaks between the proton signal at δ 4.98 (H-1_{glc}) and the carbon resonance at δ 89.4 (C-3), and between the proton signal at δ 5.25 (H-1_{gal}) and the

carbon resonance at δ 84.6 (C-3 $_{glc}$). Thus, compound **2** was defined as urs-12-ene-3β,6β,16β-triol-3-*O*-β-galactopyranosyl-(1→2)- β-glucopyranoside.

Compound **3** exhibited a quasi-molecular ion peak at *m/z* 811.4463 $[M–H]^-$ in the HR-ESI-MS-spectrum (negative mode), in accordance with an empirical molecular formula of $C_{42}H_{68}O_{15}$. The negative ESI-MS spectrum showed a psuedomolecular ion peak at *m/z* 811 $[M–H]^-$ indicating a molecular weight of 812 for compound **3**. A detailed analysis of the NMR data (^1H, ^{13}C, 2D-TOCSY, DQF-COSY, HSQC) of compound **3** in comparison with those of **1** showed a difference in the replacement of the 23-CHO function in **1** by a 23-COOH group in **3**. In particular the HMBC correlations between the proton signal at δ 1.52 (Me-24) and the carbon resonances at δ 53.6 (C-5), 55.3 (C-4), 87.1 (C-3) and 183.1 (C-23) supported the occurrence of a 23-COOH group (Mahato and Kundu, 1994). From these data, the structure of **3** was determined as 3β,6β,16β-trihydroxyolean-12-en-23-oic acid-3-*O*-β-glucopyranosyl-16-*O*-β-glucopyranoside.

The negative HR-ESI-MS of compound **4** gave a psuedomolecular ion peak at *m/z* 635.4173 $[M–H]^-$, in agreement with a molecular formula of $C_{36}H_{60}O_9$. Compound **4** displayed a molecular peak at *m/z* 635 $[M–H]^-$ in ESI-MS negative mode indicating a molecular weight of 636, as well as a fragment ion of *m/z* 473 $[M–H–162]^-$ corresponding to the loss of a hexosyl moiety. Compound **4** displayed 36 carbon resonances in its ^{13}C NMR spectrum. For the aglycone portion along with eight methyl resonances (δ 16.8, 17.4, 18.2, 18.7, 18.8, 22.9 and 25.0) and two sp^2-hybridized carbon signals (δ 126.0 and 138.6), signals for four hydroxymethine groups (δ 67.0, 67.3, 70.5 and 89.3) could be observed. A detailed analysis of the NMR data (^1H, ^{13}C, 2D-TOCSY, DQF-COSY,

HSQC) of the aglycone moiety of compound **4** in comparison with those of **2** showed differences in the carbon resonances of ring E. In the HMBC spectrum the methyl signal at δ 1.41 (Me-30) showed a long-range correlation with the carbon resonance at δ 70.5, allowing us to deduce the occurrence of an additional secondary alcoholic function at C-21 (Mimaki et al., 2004). The coupling constants of H-21 (δ 3.97, dd, J = 11.0, 4.0 Hz) and the NOESY correlation between H-21 and Me-29 suggested the β-orientation of the C-21 hydroxyl group (Mimaki et al., 2004). The sugar unit, identified as β-glucopyranosyl, was placed at C-3 of the aglycone on the basis of the HMBC correlation between the anomeric proton at δ 5.01 and the carbon resonance at δ 89.3 (C-3). Thus, compound **4** was defined as urs-12-ene-3β,6β,16β,21β-tetraol-3-*O*-β-glucopyranoside.

Compound **5** showed the same molecular formula as compound **4** ($C_{36}H_{60}O_9$), indicated by both HR-ESI-MS (*m/z* 635.4163 $[M–H]^-$) and ESI-MS analysis (*m/z* 635 $[M–H]^-$). Eight tertiary methyl signals (δ 1.23, 1.26, 1.29, 1.44, 1.54, 1.62, 1.64, 1.74) in the ^{1}H NMR spectrum along with resonances of ring E in the ^{13}C NMR spectrum showed that compound **5** differed from **4** only in the olean-12-ene skeleton instead of an urs-12-ene skeleton, being identical for the other structural features. Thus, compound **5** was defined as olean-12-ene-3β,6β,16β,21β-tetraol-3-*O*-β-glucopyranoside.

The molecular formula for compound **6** was established as $C_{42}H_{70}O_{15}$ by the negative-ion mode HR-ESI-MS showing a pseudo-molecular ion peak at *m/z* 813.4649 $[M–H]^-$. Compound **6** also exhibited a quasi-molecular ion peak in negative-ion mode ESI-MS at 813 $[M–H]^-$, indicating a molecular weight of 814. Data from the ^{13}C NMR spectrum of **6** suggested also in this case a polyhydroxylated-olean-12-ene glycoside

structure. The ^{13}C NMR spectrum showed 42 resonances, of which 30 were assigned to a triterpenoid moiety and 12 to the glycosyl portion. The following NMR data exhibited the structural features of olean-12-en-3β,6β,16β,21β,23-pentaol as the aglycone: seven tertiary methyl resonances (δ 0.88, 0.89, 0.97, 1.09, 1.19, 1.27 and 1.35), resonances at δ 3.55 (dd, J = 12.0, 4.0 Hz) 3.59 (dd, J = 11.0, 4.5 Hz), 4.10 (dd, J = 11.5, 4.5 Hz), 4.45 (br s) ascribable, respectively, to the 21β, 3β, 16β, 6β-protons on hydroxymethine carbons, one hydroxymethylene (δ 3.48 and 3.78, each a doublet, J = 12.0 Hz), and the resonance of H-12 at δ 5.29 (t, J = 3.5 Hz). On the basis of the HMBC experiment that provided unambiguous correlations between the methyl signal at δ 1.09 (Me-24) and the carbon resonances at δ 44.8 (C-4), 48.2 (C-5), 83.3 (C-3) and 64.6 (C-23), the primary alcoholic function was located at C-23 (Mahato and Kundu, 1994). Analysis of the NMR data of the sugar portion revealed the occurrence of two β-glucopyranosyl units linked at C-3 and C-16, as in compound **1**. Thus the structure of **6** was defined as olean-12-ene-3β,6β,16β,21β,23-pentaol-3-*O*-β-glucopyranosyl-16-*O*-β-glucopyranoside.

The negative HR-ESI-MS of compound **7** showed a molecular ion peak at *m/z* 913.5189 $[M–H]^-$, in accordance with an empirical molecular formula of $C_{47}H_{78}O_{17}$. Compound **7** displayed a molecular peak at *m/z* 913 $[M–H]^-$ in ESI-MS negative mode indicating a molecular weight of 914. The ^{13}C-NMR spectrum of compound **7** showed 47 signals, of which 30 were assigned to the aglycone and 17 to the saccharide portion. Analysis of ^{1}H and ^{13}C NMR spectra of the aglycone moiety (Table A.2; Appendix A) clearly showed that it differed from that of compound **5** for the absence of the secondary alcoholic function at C-21. The sugar portion of **1** exhibited, in the ^{1}H NMR spectrum,

three anomeric proton resonances (δ 4.32, d, J = 7.5 Hz; 4.43, d, J = 7.5 Hz; 4.48, d, J = 6.5 Hz). The structures of the oligosaccharide moieties were deduced using 1D-TOCSY and 2D NMR experiments, which indicated that two β-glucopyranosyl units and one α-arabinopyranosyl unit were present. The absence of any ^{13}C NMR glycosidation shift for one glucopyranosyl residue (δ 4.32) and the arabinopyranosyl unit suggested that these sugars were terminal units. Glycosidation shifts were observed for C-2_{glcI} (δ 83.5). Direct evidence for the sugar sequence and their linkage sites to the aglycone was derived from the results of the HMBC experiment that showed unequivocal correlations between resonances at δ 4.32 and 90.5 (H-1_{glc}-C-3), δ 4.43 and 77.2 (H-1_{glc}-C-16), and δ 4.48 and 83.5 (H-1_{ara}-C-2_{glcI}). Thus, compound **7** was identified as olean-12-ene-3β,6β,16β-triol-3-*O*-β-glucopyranosyl-16-*O*-α-arabinopyranosyl-(1→2)- β-glucopyranoside.

The HR-ESI-MS spectrum (negative-ion mode) of compound **8** exhibited a pseudomolecular ion peak at *m/z* 929.5114 $[M–H]^-$, ascribable to a molecular formula of $C_{47}H_{77}O_{18}$. The ESI-MS experiment (negative-ion mode) also showed a quasi-molecular ion peak at *m/z* 929 $[M–H]^-$ indicating a molecular weight of 796. Analysis of ^{1}H and ^{13}C NMR spectra of the aglycone moiety of compound **8** clearly showed that it differed from that of compound **6** for the absence of the secondary alcoholic function at C-21. Analysis of the glycosidic NMR data demonstrated that they were superimposable with those of **7**. Thus, compound **8** was defined as olean-12-ene-3β,6β,16β,23-tetraol-3-*O*-β-glucopyranosyl-16-*O*-α-arabinopyranosyl-(1->2)- β-glucopyranoside.

The molecular formula for compound **9** was established as $C_{36}H_{58}O_{10}$ by the negative-ion mode HR-ESI-MS spectrum showing a pseudo-molecular ion peak at *m/z*

649.3945 $[M–H]^-$. Compound **9** also exhibited a quasi-molecular ion peak in negative-ion mode ESI-MS at 649 $[M–H]^-$, indicating a molecular weight of 650. Analysis of NMR data of compound **9** in comparison with those of **5** clearly indicated that **9** differed from **5** only for the replacement of the 23-Me group with an aldehydic function. On the basis of the foregoing data the structure 3β,6β,16β,21β-tetrahydroxyolean-12-en-23-al-3-*O*-β-glucopyranoside was assigned to compound **9**.

To the best of our knowledge, the aglycones of compounds **1-9** are reported here for the first time. Most unusual is the co-occurrence of 6β-OH and 16β-OH functions in an olean-12-ene or an urs-12-ene series. The only other compound reported in the literature from this rare class of 6β-hydroxylated β–amryin derivatives is daturadiol, isolated first from *Datura innoxia* (Kocor and St. Pyrek, 1973), and more recently from *Terminalia brasiliensis* (Araujo and Chaves, 2005). However, daturadiol does not possess a 16β-OH.

Several studies have demonstrated that polyhydroxylated triterpenes isolated from members of the Asteraceae exhibit potent anti-inflammatory effects, as well as, anti-tumor and cytotoxic activities (Akihisa et al., 1996; Ukiya et al., 2001; Ukiya et al., 2002; Neukirch et al., 2005). Moreover, recent studies on an oleanane triterpenoid possessing both 3-β and 6-α hydroxyl groups have exhibited anti-tumor activity against a diverse panel of tumor cell lines (Wang et al., 2006). Taking into account the unusual structural features of *Silphium* saponins to determine their possible functional role as cancer preventives, we tested their effect on breast cancer cell proliferation. Initially, we tested the effect of 20 μg/ml of compounds **1-9** on growth of MDA-MB-231 human breast

cancer cells for 96 h with fresh compounds added every 48 h, as shown in Figure 5.5. At the concentration tested, only one of the saponins, compound **2** (urs-12-ene-3β,6β,16β-triol-3-*O*-β-galactopyranosyl-(1→2)- β-glucopyranoside) reduced cell number compared to the vehicle alone (DMSO).

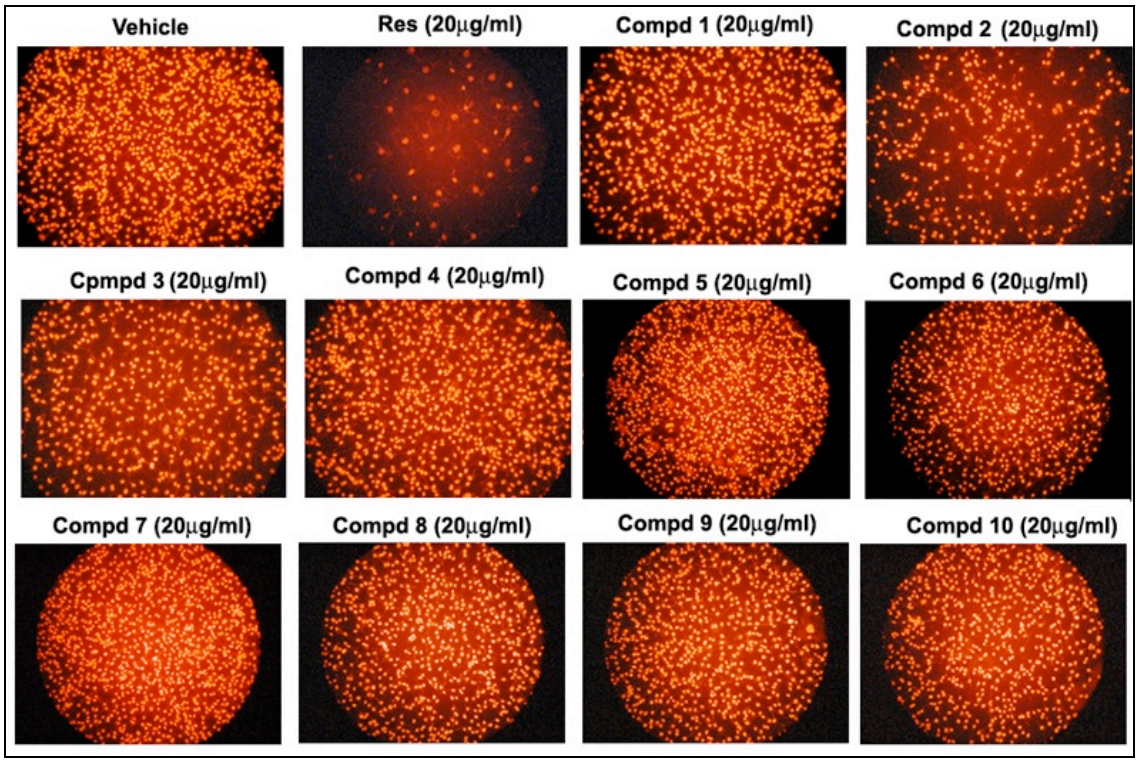


Figure 5.5. Initial testing of the effect of 20 μg/ml of compounds 1-9 on growth of MDA-MB-231 human breast cancer cells for 96 h with fresh compounds added every 48 h.

The effective concentration of compound **2** was determined in Fig. 5.6a and 5.6b, where we show that compared to the vehicle control, there was an 80% reduction in cell number at 25 μg/ml. This effect was almost 100% at 50 and 100 μg/ml. Since the observed decrease in cell number appears to be due to decreased cell proliferation (i.e. reduced number of viable nuclei) and not due to increased membrane permeability, saponins may induce inhibitory effects in cell cycle progression. This result indicates *Silphium* as a potential source for anti-breast cancer compounds.

Compound **2** was the only saponin tested in this study that possesses a galactopyranosyl residue, suggesting that this sugar residue may be involved in the cytotoxic activity. Structure-activity studies of saponins have demonstrated that the nature of the sugar moiety is very important for cytotoxic activity (Bader et al., 1996). Compound **2** also differs from all other saponins tested here by the presence of a monodesmosidic disaccharide chain attached to the C-3 position, while all other saponins tested are either bidesmosidic or monodesmosidic with a single glucopyranosyl residue attached to the C-3 position, indicating that the length and linkage of the sugar chain may also influence cytotoxicity. A survey of 16 triterpenoid saponins with different sugar linkages found monodesmosides more active than bidesmosides in killing cancer cells (Quetin-Leclercq et al., 1992).

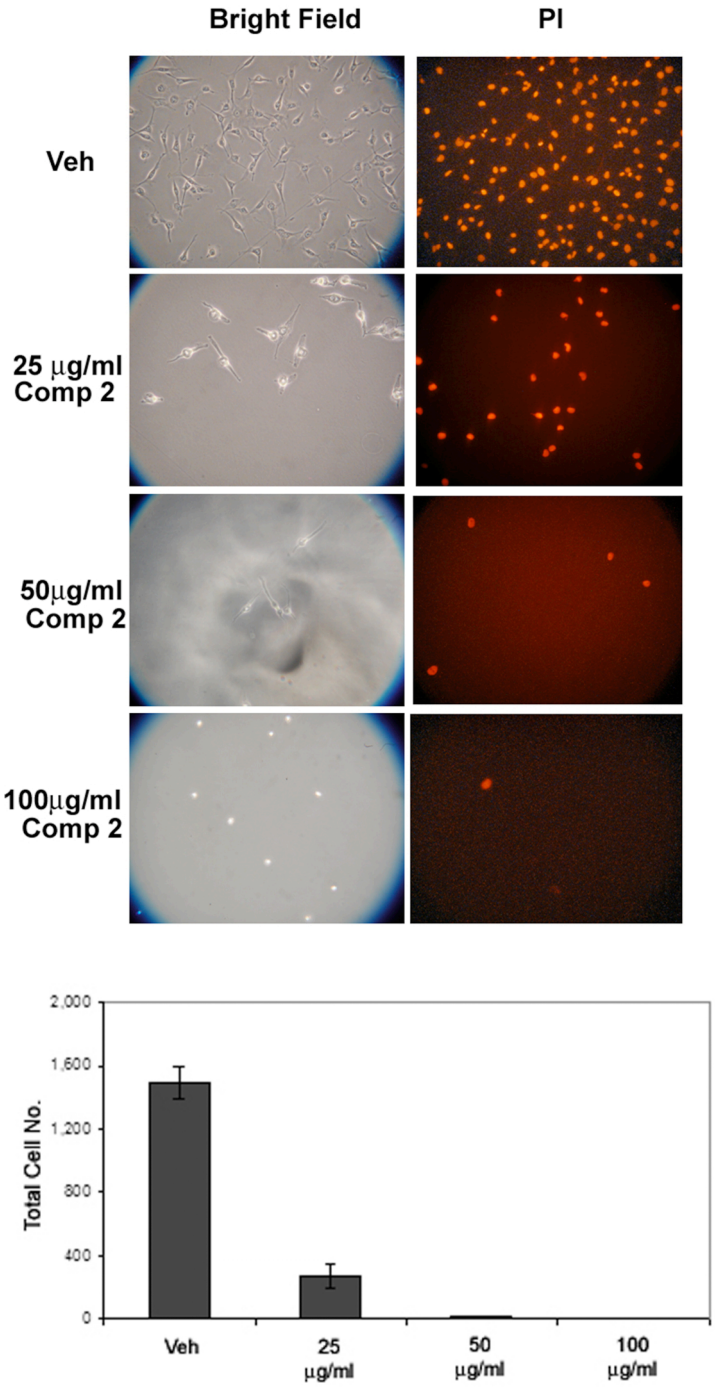


Figure 5.6a and 5.6b. Effect of Compound 2 on breast cancer cell proliferation. Semi-confluent MDA-MB-231 cells were treated with vehicle (0.1% DMSO) or compound 2 at 25, 50, or 100 µg/ml every 48 hours for 96 hours. A. Representative cells from bright field (left) or fluorescence microscopy following propidium iodide (PI) staining to visualize nuclei (right). B. Total number of cells with intact, viable nuclei following vehicle or compound 2 at 25, 50, or 100 µg/ml. Cell number was quantified from 30 microscopic fields/well. Error bars are standard deviations for N=3.

Triterpenoid saponins from other plant sources have also been shown to activate estrogen receptors (Lee et al., 2003) as well as inhibit breast cancer cell proliferation via suppression of cell cycle progression and cell survival signaling and promotion of apoptosis (Mujoo et al., 2001; Chen et al., 2003). Future studies will focus on the exact mechanism by which compound **2** exerts its anti-proliferative effects on breast cancer cells.

5.1.3 Experimental

General experimental procedure- Optical rotations were measured in MeOH with a Perkin-Elmer 241 polarimeter equipped with a sodium lamp (589 nm) and a 1dm microcell. ESI-MS was performed on a Finnigan LCQ Ion Trap mass spectrometer and HR-ESI-MS experiments were performed on Fourier Transform Ion Cyclotron mass spectrometer (Ion Spec, Varian). The spectra were recorded by infusion into the ESI (Electrospray Ionization) source using MeOH as a solvent. ^{1}H and ^{13}C NMR spectra were recorded in either pydridine-d_5 or CD_3OD on a Varian Inova 500 Mz spectrometer. All chemical shifts (δ) are given in ppm units with reference to TMS as an internal standard and the coupling constants (J) are given in Hz.

Column chromatography was carried out on Kieselgel 60 (60-200µm, Merck) or LiChroprep RP-18 (40-60µm, Merck). MPLC was performed on a Beckman 120B pump using Millipore Vantage-L (16 x 250, 32 x 250 mm) columns with a flow rate of 2ml min^{-1}. Fractions were monitored by TLC on silica gel plates (Merck precoated silica gel

60 F_{254}) and developed in the solvent system, EtOAc: HOAc:H_2O (9:2:2). Spots were visualized by heating after spraying with 10% H_2SO_4.

Plant material- Aerial parts of *S. radula* were collected by Dr. Jeffery Williams in Lee County, Texas in July 2002 near a rest area on Hwy 290 west of Giddings, Texas. Samples of *S. radula* were identified by Prof. Mark Bierner and a voucher specimen was deposited at the University of Texas at Austin Plant Resource Center (No. JW2002/7/3).

Extraction and isolation- The dried mixture of leaves and stems of *S. radula* (500 g) was successively extracted for 48 hours with CH_2Cl_2-MeOH (1:1), MeOH and MeOH-H_2O (1:1) at room temperature to give three separate crude residues. The conc. MeOH and MeOH-H_2O extracts were combined and dried under vacuum and the concentrate (70 g) was dissolved in water and loaded onto a 7.5 x10 cm C-18 column. The column was washed with water to remove sugars and then with 40% MeOH to remove phenolics. Saponins were removed with 85% MeOH and the crude saponin fraction was subjected to silica gel CC eluted with a gradient of $CHCl_3$: MeOH: H_2O (9:1:0.1 to 6:4:0.4) to give seven major fractions (I-VII). Fraction I was subjected to reversed phase C-18 MPLC with a gradient of CH_3CN:H_2O with 1% H_3PO_4 affording compound **1** (128. 6mg). Fraction II was subjected to reversed phase C-18 MPLC using a MeOH:H_2O gradient giving compound **2** (23.4 mg). Remaining saponin mixtures III-VII obtained from the above processing were repeatedly subjected to MPLC purification on silica RP-18 using a

MeOH-H_2O gradient, affording compounds **3** (15.6 mg)**, 4** (2.7 mg), **5** (4.6 mg), **6** (115.9 mg), **7** (42.9 mg), **8** (263 mg) and **9** (96.4 mg).

*3β,6β,16β-Trihydroxyolean-12-en-23-al-3-O-β-glucopyranosyl-16-O-β-glucopyranoside **(1)*** White powder; $[\alpha]_D^{20}$ +23.46 (MeOH; *c* 0.24); ^1H and ^{13}C NMR see Table A.1;Appendix A; ESI-MS (negative ion mode): *m/z* 795 $[M-H]^-$; ESI-MS-MS (795) *m/z* 633 $[M-H-162]^-$, 472 $[M-H-162-162]^-$;HR-ESI-MS (negative ion mode) *m/z* 795.45363 (calcd. For $C_{42}H_{70}O_{14}$)

*Urs-12-ene-3β,6β,16β-triol-3-O-β-galactopyranosyl-(1→2)-β-glucopyranoside **(2)*** White powder; $[\alpha]_D^{20}$ +21.78 (MeOH; *c* 0.24); ^1H and ^{13}C NMR see Table A.1;Appendix A; ESI-MS (negative ion mode): *m/z* 781 $[M-H]^-$; ESI-MS-MS (781) *m/z* 619 $[M-H-162]^-$, 457 $[M-H-162-162]^-$;HR-ESI-MS (negative ion mode) *m/z* 781.47437 (calcd. For $C_{42}H_{70}O_{13}$)

*3β,6β,16β-Trihydroxyolean-12-en-23-oic acid-3-O-β-glucopyranosyl-16-O-β-glucopyranoside **(3)*** White powder; $[\alpha]_D^{20}$ +12.67 (MeOH; *c* 0.26); ^1H and ^{13}C NMR see Table A.2; Appendix A; ESI-MS (negative ion mode): *m/z* 811 $[M-H]^-$; HR-ESI-MS (negative ion mode) *m/z* 811.44855 (calcd. For $C_{42}H_{68}O_{15}$)

*Urs-12-ene-3β,6β,16β,21β-tetraol-3-O-β-glucopyranoside **(4)*** White powder; $[\alpha]_D^{20}$ +10.90 (MeOH; *c* 0.35); ^1H and ^{13}C NMR see Table A.1;Appendix A; ESI-MS (negative

ion mode) 635 $[M- H]^-$: *m/z*; HR-ESI-MS (negative ion mode) *m/z* 635.41646 (calcd. For $C_{36}H_{60}O_9$)

*Olean-12-ene-3β,6β,16β,21β-tetraol-3-O-β-glucopyranoside (**5**)* White powder; $[\alpha]_D^{20}$ +7.88 (MeOH; *c* 0.24); ^1H and ^{13}C NMR see Table A.1; Appendix A; ESI-MS (negative ion mode): *m/z* 635 $[M- H]^-$; HR-ESI-MS (negative ion mode) *m/z* 635.41646 (calcd. For $C_{36}H_{60}O_9$)

*Olean-12-ene-3β,6β,16β,21β,23-pentaol-3-O-β-glucopyranosyl-16-O-β-glucopyranoside (**6**)* White powder; $[\alpha]_D^{20}$ +16.76 (MeOH; *c* 0.28); ^1H and ^{13}C NMR see Table A.2; Appendix A; ESI-MS (negative ion mode): *m/z* 813 $[M- H]^-$; HR-ESI-MS (negative ion mode) *m/z* 813.46420 (calcd. For $C_{42}H_{70}O_{15}$)

Olean-12-ene-3β,6β,16β-triol-3-O-β-glucopyranosyl-16-O-α-arabinopyranosyl-(1→2)-β-glucopyranoside (7) White powder; $[\alpha]_D^{20}$ +14.28 (MeOH; *c* 0.27); ^1H and ^{13}C NMR see Table A.2; Appendix A; ESI-MS (negative ion mode): *m/z* 913 $[M- H]^-$; HR-ESI-MS (negative ion mode) *m/z* 913.51663 (calcd. For $C_{47}H_{78}O_{17}$)

*Olean-12-ene-3β,6β,16β,23-tetraol-3-O-β-glucopyranosyl-16-O-α-arabinopyranosyl-(1→2)-β-glucopyranoside (**8**)* White powder; $[\alpha]_D^{20}$ +17.74 (MeOH; *c* 0.28); ^1H and ^{13}C NMR see Table A.2; Appendix A; ESI-MS (negative ion mode): *m/z* 929 $[M- H]^-$; HR-ESI-MS (negative ion mode) *m/z* 929.51154 (calcd. For $C_{47}H_{78}O_{18}$)

*3β,6β,16β,21β-Tetrahydroxyolean-12-en-23-al-3-O-β-glucopyranoside (**9**)* White powder; $[\alpha]_D^{20}$ +10.26 (MeOH; *c* 0.23); ^1H and ^{13}C NMR see Table A.2; Appendix A; ESI-MS (negative ion mode): *m/z* 649 $[M- H]^-$; HR-ESI-MS (negative ion mode) *m/z* 649.39572 (calcd. For $C_{36}H_{58}O_{10}$)

Biological assays- Cell culture. MDA-MB-231 human breast cancer cells (from the American Type Culture Collection (ATCC)) were cultured in a Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% fetal bovine serum (FBS) at 37.9702 °C in 5% CO_2.

Cell proliferation. Semi-confluent cells seeded on 6 well plates in phenol-red free DMEM and 5% FBS (charcoal-stripped) were treated with vehicle or compounds **1-9** every 48 hours for 96 hours. Stock solutions were in 100% DMSO and were diluted in DMEM to a final concentration of 0.1% (v/v) DMSO prior to treatment. Cells were fixed in methanol, stained with propidium iodide (PI), and intact nuclei were counted from digital images acquired by an Olympus CKX41 microscope. Total number of cells was quantified from 30 microscopic fields/well from experiments carried out in triplicate.

5.2. The Triterpene Saponin Chemistry of *Silphium terebinthinaceum* and *S. laciniatum*

5.2.1 Introduction

Silphium terebinthinaceum and *S. laciniatum* stand out among *Silphium* species in that they produce unusually tall flowering stems (up to 4.5 m tall) that harbor diverse and well-documented insect communities (Weaver, 1954; Tooker and Hanks, 2002, other Tooker publications). Moreover, *S. terebinthinaceum* and *S. laciniatum* are characteristic species of tallgrass prairies, an ecosystem threatened by extreme habitat loss due to agriculture and commercial developments. Tallgrass prairies once covered over 170 million acres, extending from Indiana to the Rocky Mountains. Today, less than 4% remains, making the prairie the most endangered ecosystem in North America (Sampson and Knopf, 1994).

Despite extensive ethnobotanical data indicating the medicinal uses of *S. laciniatum* and *S. terebinthinaceum* by Native American tribes (see Chapter 2), information on the chemistry of the two species is limited to a few publications on monoterpene volatiles involved in tritrophic plant/insect interactions (Tooker et al., 2002; Tooker et al., 2005), and a doctoral dissertation reporting on several ubiquitous phenolic compounds (Williams, 2006). To the best of our knowledge, there is currently no literature data on triterpene compounds from either *S. terebinthinaceum* or *S. laciniatum*, however there is an unpublished account of pentacyclic triterpenes isolated from *S. laciniatum* that exhibit potent anti-tuberculosis activity, showing >90% inhibition at

100μg/ml (personal communication, Dr. Kirk Manfredi, Prof of Chemistry, University of Northern Iowa). Moreover, biological screening of organic extracts of the leaves, stems and roots *S. laciniatum* revealed moderate activity in an anti-cancer screening (LC_{50} > or = 20% of the cell lines) and aqueous extracts of *S. laciniatum* showed strong anti-HIV activity (LC_{50} > 50% compared with the control) (Kindscher et al., 1998).

S. terebinthinaceum var. *pinnatifidum* has been suggested to be a hybrid between *S. laciniatum* and *S. terebinthinaceum* because of their similarity in leaf shape. Hybrid specimens have been documented in numerous localities in Illinois and Ohio (Settle and Fisher, 1972; personal communications, Dr. Ken Roberston and Dr. Jennifer Clevinger) (Figure 5.7). However, both morphological and molecular studies suggest that the hybrids tested were more closely allied to S. *terebinthinaceum*. Still, controversy remains about whether any of the proposed "hybrid populations" represent true genetic crosses between *S. laciniatum* and *S. terebinthinaceum*.

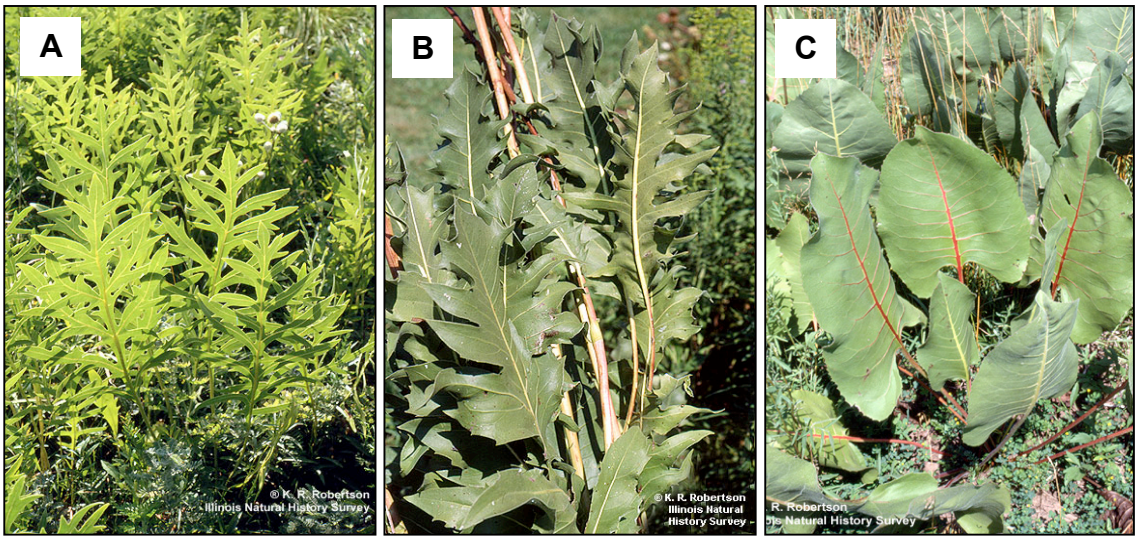


Figure 5.7. Photographs of leaves of A) *S. laciniatum* B) possible hybrid between *S. laciniatum* and *S. terebinthinaceum* and C) *S. terebinthinaceum*. Photographs provided courtesy of Kenneth Robertson.

Triterpene saponins have been reported from several *Silphium* species (Davidyants et al., 1984a,b; Davidyants et al., 1985; Davidyants et al., 1986; Calabria et al., 2008; Kowalski, 2007). Our preliminary chemical screening of all *Silphium* species indicated the presence of triterpene saponins in the leaves of *S. laciniatum* and *S. terebinthinaceum*. A more detailed, population level screening of intraspecific variation of saponins in *S. terebinthinaceum* and *S. laciniatum* should help to clarify the taxonomic position of suspected hybrid specimens. Thus, triterpene saponins from the leaves of *S. terebinthinaceum*, *S. laciniatum* and potential hybrids were extracted, purified and chemically characterized using thin layer chromatography (TLC) and reverse phase HPLC with electrospray ionization mass spectrometry (ESI/MS). In addition, roots of the two species were also examined for saponin content by the same methods.

5.2.2 Methods

Plant material- Leaves of both *S. terebinthinaceum* and *S. laciniatum* were collected from five tallgrass prairies in Illinois by Dr. Kenneth Robertson (curator and plant systematist at the Illinois Natural History Survey). 10-15 leaves from each species at each collection site were harvested at the beginning of flowering, during the first week of June, when the concentrations of saponins in the leaves are reported to be at their highest (Kowalski, 2007). Leaves were selected from different locations throughout each prairie, not from plants growing next to each other. The collection sites were located at: Prospect Cemetery Prairie Nature Preserve, Ford County, Illinois; Loda Cemetery Prairie Nature Preserve, Iroquois County, Illinois; Weston Cemetery Prairie Nature Preserve, McLean

County, Illinois; Broughton Cemetery Prairie, Livingston County, Illinois; and the Don Gardner Prairie Project. Ford County, Illinois. The samples were pressed flat, placed in a plant press and air-dried.

Additional leaves were collected from plants that displayed morphological characteristics of hybrids between *S. terebinthinaceum* and *S. laciniatum*. These specimens were collected from the Broughton Cemetery Prairie, the Don Gardner prairie and two additional sites, the Meadowbrook Park Prairie and Barhnet Prairie Nature Preserve, both in Champaign County, Illinois. Roots of *S. terebinthinaceum* and *S. laciniatum* were harvested from the Don Gardner Prairie Restoration site on April 9, 2007 by Don Gardner (Gardner Prairie Project).

Extraction and Isolation- Dried leaves (2g) of *S. terebinthinaceum*, *S. laciniatum* and proposed hybrids were extracted 2X with 85% MeOH and the resultant extracts were filtered and dried under vacuum. The crude methanolic extract of each species was resuspended in H_2O and passed through a C-18 column preconditioned with H_2O. The column was washed with H_2O to remove sugars and then with 40% MeOH to remove phenolics. Saponins were eluted with 85% MeOH and the solution was evaporated and the residue weighed. Fractions were monitored by TLC on silica gel plates (Merck precoated silica gel 60 F_{254}) and developed in the solvent system, EtOAc: HOAc:H_2O (9:2:2). Spots were visualized by heating after spraying with 10% H_2SO_4.

High performance liquid chromatography- HPLC analyses were performed on a Thermo Separation Products HPLC system coupled with a photodiode array (PDA) detector and a Thermo Finnigan (San Jose, CA) LTQ XL Linear Ion Trap Mass Spectrometer in sequence. A Thermo Hypersil GOLD column (50mm x 2.1mm, 3µm) was used for separation using a mobile phase gradient consisting of water (A) and acetonitrile (B). Formic acid (0.1%) was used as a mobile phase modifier to facilitate ionization and to achieve a high level of chromatographic separation efficiency and ESI-MS instrument response. The system was run with the following gradient program: from 5-95%B in 2 minutes, equilibration 2 minutes, or from 5-95%B in 25 minutes, from 95-100%B in 2 minutes and held for 4 minutes before returning to 5%B in 3 minutes. A flow was set at 0.2 mL/min and directed to the ion trap mass spectrometer via an ESI interface. The mass spectrometer was operated in both positive and negative modes. A sample injection volume of 10µl was used for all analyses performed.

Mass Spectrometry- Negative and positive ion ESI-MS experiments were performed using a Thermo Finnigan (San Jose, CA) LTQ XL Linear Ion Trap Mass Spectrometer equipped with an electrospray ion source. The ESI conditions were as follows: spray voltage 5.0 kV; capillary offset voltage -35 V; capillary temperature 300.00 °C. Mass spectra were recorded in the range of 200.00-2000.00 *m/z*.

5.2.3 Results

TLC analysis of leaves from *S. terebinthinaceum* and *S. laciniatum*

The results from our preliminary TLC screening of leaf extracts of *S. terebinthinaceum* and *S. laciniatum* indicated that the concentration of saponins in individual leaf samples (2g) was below the limits of detection for TLC. A few spots corresponding to polar flavonoids were detected in the leaf extracts, which is in good agreement with previous chemical data for these species (Williams, 2006). The flavonoid patterns observed in our preliminary TLC screening of *S. terebinthinaceum* and *S. laciniatum* were qualitatively very similar for leaves collected from the five different field sites and between individuals in the same field site, leading to the conclusion that there is a low level of variability in flavonoid chemistry between the two species and in the field sites tested.

HPLC/ESI-MS analysis of leaves from *S. terebinthinaceum* and *S. laciniatum*

Five extracts of *S. terebinthinaceum and S. laciniatum* were submitted for initial LC-MS screening, consisting of a single representative sample for each species and three hybrid samples. There were no saponins detected in the LC-MS analysis of the *S. terebinthinaceum* and *S. laciniatum* leaf extracts or in the hybrids. These results were not congruent with our comprehensive LC-MS analysis of all *Silphium* species (see discussion in Chapter 6), which confirmed the presence of a few saponins in the leaves of both species, but at very low concentrations in comparison with all other species of *Silphium*. In addition, earlier TLC screening of methanolic extracts from 600g of leaf

material, also confirmed the presence of saponins. However, the quantity tested was over 100x greater than the quantity tested in the current study (2g samples).

In order to eliminate instrumentation error as a cause for the negative results obtained in the current analysis, a series of LC-MS experiments were performed. First, to address the possibility of errors introduced during sample preparation, different extracts of *S. terebinthinaceum* and *S. laciniatum* that were previously confirmed to contain small concentrations of saponins (see Chapter 6) were retested under the LC-MS conditions used in our current analysis. The results from this test were inconclusive. Although there were peaks detected with the molecular weights and retention times expected for saponins, there were no clear fragmentation ions corresponding to the aglycones, loss of sugars or dimers. Moreover, the observed molecular ions did not correspond to the saponins previously analyzed and confirmed in our earlier analysis of the same extracts. Therefore, instrumentation differences are probably responsible for the differing results obtained in our current studies.

An additional test was conducted with a pure saponin standard previously isolated from *Silphium radula* to ensure that saponins within the concentration range expected for *S. terebinthinaceum* and *S. laciniatum* could be detected using the current method. The saponin standard was successfully detected, however, the results demonstrated that differences in the acquisition parameters used such as spray voltage, capillary offset voltage and capillary temperature greatly effect the ionization of the saponins being examined and it is likely that the current instrument settings were not sufficient for characterizing saponins in this concentration range.

It is not possible to determine whether the concentration of saponins in the leaf extracts tested here is extremely low or completely absent. Still, the apparent lack of saponins in the leaf extracts of *S. terebinthinaceum* or *S. laciniatum* in comparison with all other *Silphium* species led us to question if the two species might store saponins in an organ other than leaves, possibly in their unusually large taproots. To test this hypothesis, roots of both species were collected from their native tallgrass prairie habitat in Illinois and examined using the same TLC and HPLC-MS methods described here for the leaves.

TLC and HPLC-MS analysis of taproots from *S. terebinthinaceum* and *S. laciniatum*

Figure 5.8 shows the taproots of *S. terebinthinaceum* and *S. laciniatum* harvested from the Gardner Prairie, Illinois. TLC results for both *S. terebinthinaceum* and *S. laciniatum* roots confirmed the presence of several triterpene saponins present in relatively high concentrations. The two species exhibited similar saponin profiles based on the Rf values observed on developed TLC plates.

LC-MS analysis of the root extracts *S. terebinthinaceum* and *S. laciniatum* were conducted to gain more information on the saponin composition and distribution in the two species. Negative mode was employed for increased sensitivity due to stronger ionization of the analytes. In addition the presence of a strong molecular ion and the typical dimer observed with this ionization mode also helped to support identifications of individual saponins. Positive mode was additionally used to confirm the negative ionization results and to gain additional information on fragmentation behavior of the

saponins. Table 5.1 shows the LC-MS results from the analysis of roots of *S. terebinthinaceum* and *S. laciniatum.*

Figure 5.8. The taproots of *S. laciniatum* (right) and *S. terebinthinaceum*. Photographs courtesy of Kenneth Robertson.

The negative mode LC-MS analysis of *S. laciniatum* and *S. terebinthinaceum* root extracts confirmed that the two species possess qualitatively similar saponin profiles. A strong molecular ion and a dimer were present for most compounds. In some cases an adduct ion (+ *m/z* 46), corresponding to an addition of formic acid was observed. There were few fragment ions present in negative mode, with the exception of compound **4**, which showed the sequential loss of a pentose $[M-132-H]^-$ and a glucuronic acid moiety $[M-132-176-H]^-$.

Figure 5.9 shows the LC chromatograms and ESI-MS spectra for analysis of *S. laciniatum* and *S. terebinthinaceum* root extracts. Although, the concentration of saponins differs between the two species, the saponin composition is identical (Table 5.1). The MS spectrum of compound 3 shown in Figure 5.9 provides an example of the identical mass fingerprints observed in the MS analysis of *S. laciniatum* and *S. terebinthinaceum* root extracts.

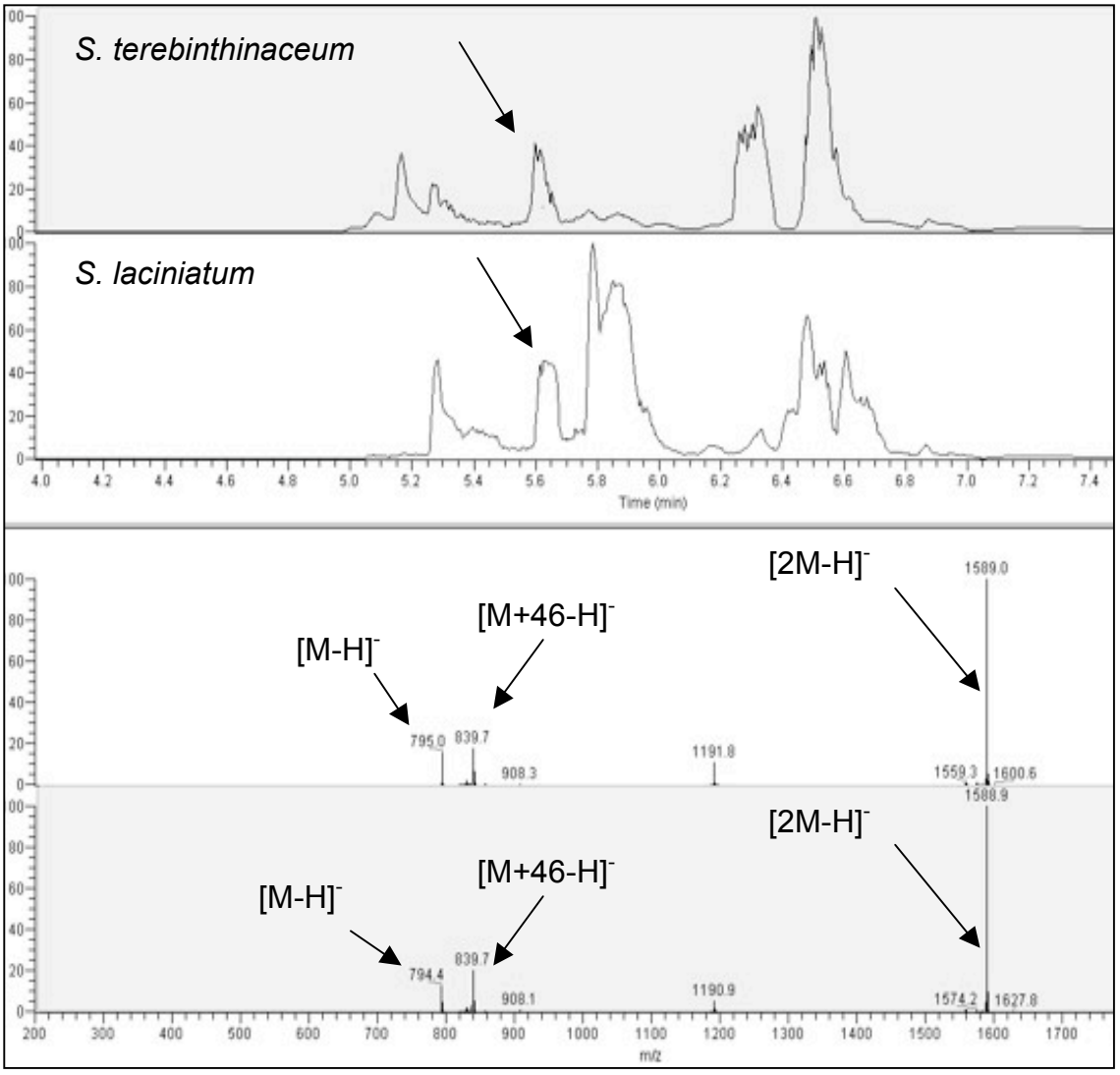


Figure 5.9. LC chromatograms and ESI-MS spectra of *S. laciniatum* and *S. terebinthinaceum* root extracts.

Peak	Time	TERE	LACI	Negative Ion mode
1	5.28	x	x	957 [M-H]-, 1913 [2M-H]-
2	5.4	x	x	927 [M-H]-, 1853 [2M-H]-
3	5.62	x	x	795 [M-H]-, 841 [M+46-H]-, 1589 [2M-H]-
4	5.72	x	x	765 [M-H]-, 811 [M+46-H]-, 633 [M-132-H], 455 M-176-H]-, 1263 [2M-H]-
5	6.32	x	x	989 [M+46-H]-, 943 [M-H]-
6	6.6	x	x	941 [M-H]-

Table 5.1. ESI-MS results for the analysis of saponins from roots of *S. laciniatum* (LACI) and *S. terebinthinaceum* (TERE).

The positive mode LC-MS analysis of *S. laciniatum* and *S. terebinthinaceum* root extracts supported the results of the negative mode LC-MS analysis and provided additional fragmentation information. However, because positive mode is less sensitive, not all compounds were detected.

5.2.4 Discussion and Future Work

This is the first LC-MS report confirming the presence of triterpene saponins in *S. terebinthinaceum* and *S. laciniatum* roots. To our knowledge, there is no literature data on triterpene saponins in flowers of either species. This would be a recommended area for future research, as a few reports have confirmed that triterpenes and their glycosides occur in relatively large amounts (up to 2% of the dry mass) in flowers of other *Silphium* species, as well as other Asteraceae (Szakiel et al., 2005; Kowalski, 2007; Kirk Manfredi, personal communication). Furthermore, many Asteraceae triterpenes possess potent anti-inflammatory and anti-tumor activities (Akihisa et al., 1996; Ukiya et al., 2001; Ukiya et al, 2002).

One possible explanation for the observed differences in saponin concentration in the leaves of *S. laciniatum* and *S. terebinthinaceum* compared with all other *Silphium* species may involve the tritrophic interactions between *Silphium*, the gall-wasp, *Antistrophus* and its natural enemy the parasitoid *Eurytroma*. Several species of gall wasps (Hymenoptera: Cynipidae) induce galls on *Silphium*. Gall-inducing insects are typically specialist plant feeders with most species confined to one specific host plant species. An excellent example involves several gall-inducing wasp species of the genus *Antistrophus* whose larvae induce galls in the pith and cambium of stems of *Silphium* spp. when female gall wasps lay eggs in the bolting stems (typically during May and June) (Tooker, 2004; Tooker and Hanks, 2006). Figure 5.10 shows a longitudinal section of stems of *S. laciniatum* and *S. terebinthinaceum* revealing *Antistrophus* galls.

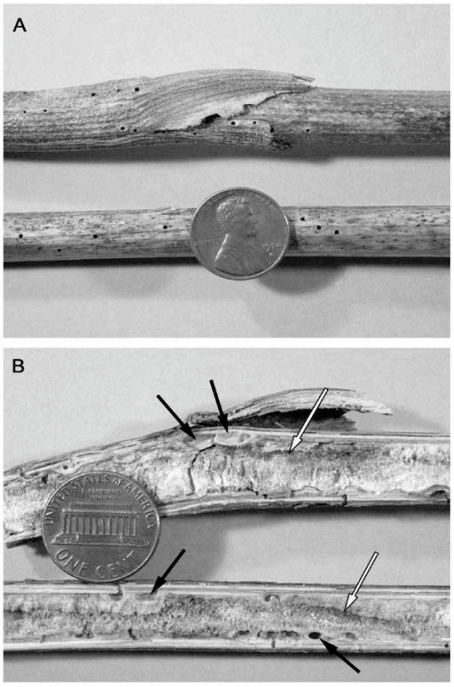


Figure 5.10. A longitudinal section of stems of *S. laciniatum* and *S. terebinthinaceum* revealing *Antistrophus* galls (photograph courtesy of John Tooker, from: Tooker, 2002).

A specific blend of monoterpenes emitted from the host plant is critical for female wasps to recognize their specific host plants (Tooker et al., 2005). In addition, upon emergence from galls, adult males use olfactory signals from their natal host plant species to locate their mates nested within the dead over-wintering stems of *Silphium.* This same blend of volatiles emitted by *Silphium* also attracts the parasitic wasp *Eurytoma lutea* Bugbee (Hymenoptera: 23 Eurytomidae), which is a natural enemy of *Antistrophus*. The larvae of *E. lutea* kill gall wasp larvae, diminishing the amount of plant damage inflicted by *Antistrophus,* thus allowing the plants to put resources into reproduction that would have otherwise been consumed by the gall wasp larvae. As a result, galled *S. laciniatum* plants that are able to attract *E. lutea* appear to have improved fitness relative to galled controlled plants (Tooker and Hanks, 2006). The molecular and biochemical basis for these interactions is largely unstudied.

The manipulation of host plants chemistry by galling (and many other) insects is well documented. Modes of action range from the induction of the hypersensitive response, suppression of genes involved in biosynthesis of defensive compounds, detoxification of active defense compounds and even tissue-specific redistribution of secondary metabolites, to name a few. For example, the nutritive tissues lining galls often are void of secondary metabolites, whereas the exterior of the galls can have relatively high concentrations of toxins. This distribution of secondary compounds gives galling insects a toxin-free food source while providing the gall insects protection against natural enemies or herbivores that might favor plant tissue (personal communication, Tooker).

It's possible that the entrance of the parasitoid *Eurytoma* into the stems is somehow facilitated by the small quantities of saponins present, thereby indirectly providing *Silphium* with a defense against *Antistrophus*. Whether or not the biosynthesis of triterpenes in these two species is altered in leaves of *S. laciniatum* and *S. terebinthinaceum* by signals from *Eurytoma, Antistrophus* or any other biotic or abiotic factor is unknown. However, the detrimental effect of some saponins on insects is well documented. For example, the tomato saponin, tomatine, has been shown to reduce the growth of insect larvae and to be generally toxic to insects (Duffey and Stout, 1996). In addition, saponins often possess potent fungicidal and anti-bacterial activities, including saponins previously isolated from *Silphium perfoliatum* (Davidyants et al., 1997). Yet, *S. laciniatum* and *S. terebinthinaceum* do not appear to require large quantities of saponins in their leaves during flowering for protection against insects or pathogens. The structure and biosynthesis of triterpenes are similar to plant and animal hormones and its possible that, like hormones, only small amounts are required to exhibit profound effects on plant growth and development.

Fire may be another possible explanation for the low concentration of saponins in leaves and the high concentrations of saponins detected in the roots of *S. terebinthinaceum* and *S. laciniatum.* Fire is a natural and necessary disturbance in healthy tallgrass prairie ecosystems; without fire, woody plants and shrubs would quickly encroach. Plants native to tallgrass prairies actually thrive from spring fires, which clear away last years plant debris, warming the soil and letting in light (Klinkenborg, 2007). Because *S. terebinthinaceum* and *S. laciniatum* are perennial species, their above ground

parts die back each year and vital nutrients and defensive compounds may be re-allocated from above ground leaves and stems to underground taproots for storage.

Furthermore, the presence of defensive compounds, such as saponins in roots of *S. terebinthinaceum* and *S. laciniatum* probably serve an allelopathic role. Studies conducted by Szakiel and coworkers (1995; 2003; 2005) have demonstrated the allelopathic effects of saponins in *Calendula officinalis* (also belonging to the Asteraceae), including their protective effects against soil pathogens and the transport of saponins from roots into the surrounding medium (Szakiel, 2003). Further studies are needed to determine the precise chemical structures of the saponins detected in roots of *S. terebinthinaceum* and *S. laciniatum.*

The present study provides new chemical data for triterpene saponins in *S. terebinthinaceum* and *S. laciniatum* leaves and roots. Because no detectable levels of saponins were found in the leaf extracts tested in our current analysis, we were unable to determine intraspecific variation for either species or provide any new insights into the chemistry of proposed hybrids between *S. terebinthinaceum* and *S. laciniatum*. Although previous reports on the production of saponins in *S. integrifolium, S. perfoliatum and S. asteriscus* (which all belong to *Silphium* section *Silphium* that are distinguished by having fibrous roots) indicate saponins are at their highest concentrations just before and during early stages of flowering (Kowalski, 2007), this was not the case for the *S. terebinthinaceum* and *S. laciniatum* tested in the present analysis. It's possible that because *S. terebinthinaceum* and *S. laciniatum* belong to section *Compositum*, which is characterized by having large taproots, that developmental distribution of saponins may

be different for species in section *Silphium*. Future work should focus on establishing a quantitative profile for saponin concentrations in leaves throughout the year to determine any differences between the two *Silphium* sections.

Once this baseline data is established, a new qualitative LC-MS method can be developed and tuned specifically for the saponins present in the leaves. If these results further support our current findings that only low, or possibly no detectable quantities of saponins are produced by *S. terebinthinaceum* and *S. laciniatum* at any time throughout the year, then future efforts should concentrate on discerning between the many environmental, developmental and genetic variables that contribute to saponin production in these species. For example, setting up a series of common garden and growth chamber experiments may be useful for identifying the exact mechanisms of control for the biosynthesis and storage of these compounds.

Furthermore, the isolation and characterization of saponins from roots of *S. terebinthinaceum* and *S. laciniatum* will be important for determining their biological functions. Preliminary anti-microbial and anti-bacterial screening of individual compounds should help to establish their potential allelopathic effects.

5.3. Isolation and Characterization of Triterpene Saponins from *Silphium integrifolium, S. morhii* and *S. perfoliatum* and Discussion of Preliminary Chemical Analysis of Triterpene Saponins from Five Other *Silphium* Species

This chapter outlines the chemical characterization and metabolic profiling of triterpene saponins from *Silphium* species that were not discussed in Chapters 5.1 or 5.2. Preliminary HPLC screenings of all *Silphium* extracts were conducted to target species that warranted further chemical investigation. *Silphium integrifolium, S. morhii,* and *S. perfoliatum,* as well as *S. asteriscus* and *S. compositum* were chosen for large-scale isolation and characterization of saponins due to the high concentrations and diversity of saponins present in their crude extracts. While the chemical analyses of the first three species were successful, difficulties were encountered in the purification of complex saponin mixtures from *S. asteriscus* and *S. compositum*. Although some partially purified saponins were obtained from these latter two species, the resulting structural data (NMR and MS) could not be easily interpreted without further purification of the mixtures, thus, HPLC/ESI-MS was employed for the tentative characterization of saponins from these species (see Chapter 6 results for details).

The species *S. albiflorum, S. brachiatum* and *S. wasiotense* were not chosen for large scale isolation due to the low concentration of saponins present in their crude extracts or as in the case of *S. albiflorum*, the lack of sufficient amounts of leaf material for large scale isolation and characterization. *S. albiflorum* is a Texas endemic with limited distribution in just two counties and large collections of leaf material (at least 300g dry weight would be necessary) would not be possible without decimating local

populations. Thus, only partial chemical characterization of triterpene saponins from these species using HPLC/ESI-MS was possible (see Chapter 6 for details).

This chapter will therefore focus on structural data for triterpene saponins from only *S. integrifolium, S. morhii* and *S. perfoliatum*. Preliminary NMR and MS data will be described for new and previously reported saponins isolated for the first time from *S. integrifolium* and *S morhii*. In addition, the biological activities of some of these compounds will be outlined. Finally, an HPLC/ESI-MS metabolic profile for saponins from *S. perfoliatum* cultivated in Poland will be discussed and compared with the data from *S. perfoliatum* collected from its native habitat in the U.S.

5.3.1 *Silphium integrifolium*[4]

Silphium integrifolium, also known as wholeleaf rosinweed, is a typical species of open rocky woodlands and prairies throughout the central United States. Several Native American tribes considered *S. integrifolium* a medicinal herb, and prepared extracts of leaves and roots for the treatment of kidney diseases, as a diuretic and an analgesic (Shemluck, 1982; Moerman, 2003). Recent molecular studies place *S. integrifolium* as a close ally to *S. perfoliatum*, a species with a well-documented chemistry including several biologically active triterpene saponins (Davidyants et al., 1984 a, b, c) (for details on biological activities, see Chapter 2). Moreover, our research group has recently reported on the isolation and structural elucidation of nine new polyhydroxylated pentacyclic oleanene and ursene-type triterpene saponins from methanolic extracts of

[4] The analyses discussed in this section were produced through the joint efforts of Sonia Piacente, Su Dharmawardane and Lalita Calabria.

leaves and stems of *Silphium radula* Nutt. (Calabria et al., 2008). The cytotoxic activities of the isolated compounds were tested against human breast cancer cell line 25 MDA-MB-231, and the results showed that saponin **I** decreased cell proliferation in a statistically significant manner at 25μg/ml.

In an effort to further characterize triterpene saponins of chemosystematic and medicinal significance from *Silphium*, our research group chemically investigated *S. integrifolium*. We report here the isolation and structural elucidation of 1 new pentacyclic ursene-type triterpene saponin (**1**) and three known oleanene-type triterpene saponins (**2-4)** from leaf and stem methanolic extracts of *Silphium integrifolium.* To our knowledge, this is the first chemical report including full spectral characterization of triterpene saponins (1D and 2D NMR and MS) for this species. Moreover, the cytotoxic activities of the isolated saponins were tested on human breast cancer line 25 MDA-MB-231, and the results showed that the compounds had no effect on breast cancer cell proliferation.

Results and Discussion

The concentrated methanolic extract of a mixture of *S. integrifolium* leaves and stems was subjected to solid phase extraction on reverse phase C-18, CC on silica gel and MPLC on reversed phase C-18, affording one new saponins (**1**) and 3 known saponins (**2-4**). Their structures were established by 1D and 2D NMR spectroscopy and ESI-MS analyses.

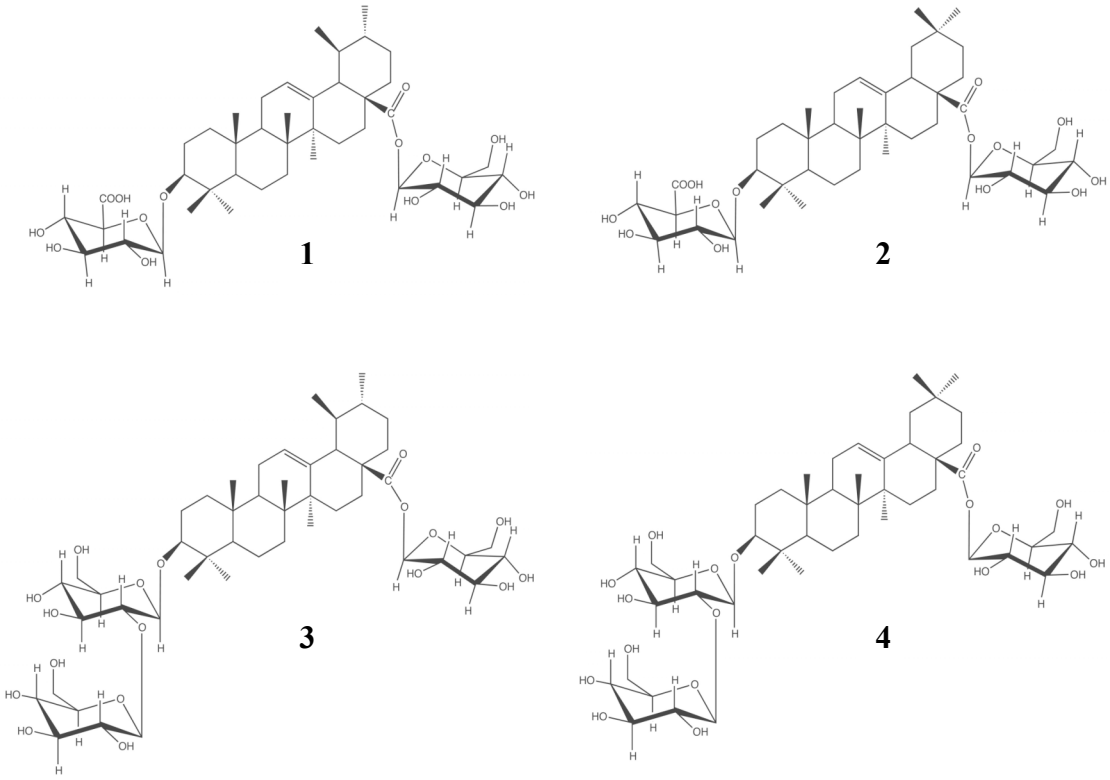


Figure 5.11. Structures of triterpene saponins isolated from *Silphium integrifolium*. Compound 1 was first isolated from *Cynara cardunculus* (Shimizu et al., 1988). Compounds 2 (Silphioside G) and 4 (Silphioside E) were previously isolated from *Silphium perfoliatum* (Davidyants et al., 1984 a, b, c). Compound 3 is a new saponin reported here for the first time.

Structures of triterpene saponins isolated from *Silphium integrifolium* are provided in Figure 5.11. See Appendix B for all NMR experiments. Compounds **1** and **2** were both obtained as amorphous white solids and were deduced to have the molecular formula $C_{42}H_{66}O_{14}$, as observed in the negative ion ESI-MS spectra at *m/z* 793. Compounds **3** and **4** were also obtained as amorphous white solids and both were deduced to have the molecular formula $C_{48}H_{78}O_{18}$ as observed in the negative ion ESI-MS spectra at *m/z* 941. Compounds **1** and **2** and compounds **3** and **4** were isolated as

isomeric mixtures and chemical shifts of all compounds were in good agreement with literature values (Shimizu et al., 1988; Davidyants et al., 1984 a, b, c).

Compound **1**, also known by the trivial name cynarasaponin C, was first isolated from *Cynara cardunculus* (Shimizu et al., 1988) and later found also in *Ilex kudincha* (Nishimura et al., 1999). The latter report found that cynarasaponin C exhibited significant acyl CoA cholesteryl acyltranferase (ACAT) inhibitory activity, which is associated with reducing inflammation and improving vascular function in hypercholesterolemia (Kharbanda et al., 2005). Compounds **2** (silphioside G) and **4** (silphioside E) were previously isolated from *Silphium perfoliatum* (Davidyants et al., 1984 a, b, c). Compound **3** is a new saponin reported here for the first time.

Taking into account the biological activities of *Silphium* saponins to determine the possible role of the saponins from *S. integrifolium* as cancer preventives, we tested their effect on breast cancer cell proliferation. Initially, we tested the effect of 20 μg/ml of compounds **1-4** on growth of MDA-MB-231 human breast cancer cells for 96 h with fresh compounds added every 48 h. At the concentration tested, only one of the saponins, compound **1** (ursolic acid 3-*O*-β-glucuronopyranosy-28-*O*-β-glucopyranoside) reduced cell number compared to the vehicle alone (DMSO).

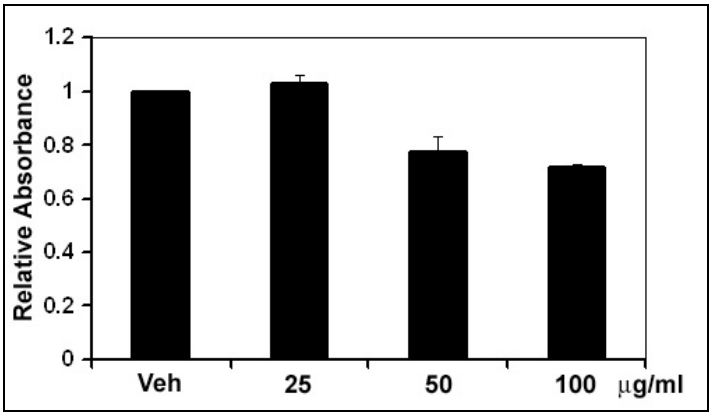


Figure 5.12. Effect of Compound 1 on breast cancer cell proliferation. Semi-confluent MDA-MB-231 cells were treated with vehicle (0.1% DMSO) or compound 1 at 25, 50, or 100µg/ml every 48 hours for 96 hours. Total number of cells with intact, viable nuclei following vehicle or compound 1 at 25, 50, or 100µg/ml. Cell number was quantified from 30 microscopic fields/well. Error bars are standard deviations for N=3.

The effective concentration of compound **1** was determined in Fig. 5.12, where we showed that compared to the vehicle control, there was no significant reduction in cell number at 25, 50 and 100µg/ml. Future biological testing on saponins isolated from *Silphium* should focus on screening the antimicrobial activities of these compounds, since many saponins are produced constitutively in plants and are thought to be involved in plant disease resistance (Papadopoulou et al., 1999).

5.3.2 *Silphium morhii*

Silphium morhii, also known as shaggy rosinweed, is an endemic species to the Appalachian Mountains of Alabama, Georgia and Tennessee and is distinguished from other *Silphium* species by its pale yellow flowers and shaggy pubescence covering the stems and leaves (Clevinger, 1999). Very little is known about the chemistry of *S. morhii*.

A dissertation report describing the LC-MS analysis of flavonoids of *S. morhii* was recently published (Williams, 2006).

In an effort to further characterize triterpene saponins of chemosystematic and medicinal significance from *Silphium*, we examined *S. morhii*. Although several saponins were partially purified from the crude methanolic leaf extract only one compound was fully characterized; the known saponin Ilexoside XLVIII, which was first isolated from *Ilex rotunda* (Amimoto et al., 1993). This is the first report of a triterpene saponin from *Silphium morhii*. The structure of Ilexoside XLVIII (compound **1**) was elucidated by NMR and mass spectrometry.

Results and Discussion

Compound **1** was obtained as an amorphous white solid and was deduced to have the molecular formula $C_{42}H_{66}O_{15}$, as observed in the negative ion ESI-MS spectra at *m/z* 809. The ^{1}H and ^{13}C-NMR spectra were in good agreement with values first reported by Amimoto et al. (1993). The structure of compound **1** isolated from *S. morhii* is provided in Figure 5.13.

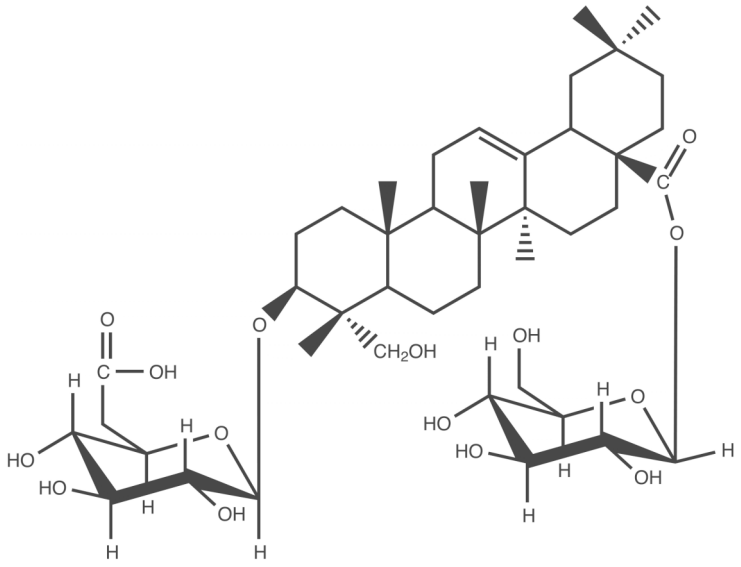


Figure 5.13. Structure of compound 1 isolated from *S. morhii*, which was first isolated from *Ilex rotunda* and is known as Ilexoside XLVIII.

Additional NMR and MS data for compound **1**, as well as data for several other saponins that were partially purified from *S. morhii* **(2-4)**, are provided in Appendix B. Further characterization of these compounds is warranted because the aglycone for compound **1**, known as hederagenin, was not previously reported from *Silphium*. Furthermore, hederagenin is extremely rare in the Asteraceae family, with only one literature report published to date (Mehta et al., 2004). Preliminary analysis of NMR data for the partially purified saponins also showed the characteristic chemical shifts for hederagenin.

5.3.3 *Silphium perfoliatum*[5]

Silphium perfoliatum was historically the first species of *Silphium* to be chemically analyzed, revealing the presence of several novel triterpene saponins (Davidyants et al., 1984 a, b, c) with a range of biological and medicinal activities that have been discussed in detail in this dissertation (Davidyants et al., 1997) (see Chapters 2 for a complete list of references). This section of Chapter 5.3 will address the topic of intraspecific variation in saponin composition within *Silphium perfoliatum*. Samples of *S. perfoliatum* cultivated in the experimental gardens at the University of Medical Sciences in Poznan, Poland will be compared with samples collected from their native habitat in Giles Co. Virginia, U.S.A. *S. perfoliatum* is an ideal model for studying intraspecific variation of saponins because, it is one of the few species of *Silphium* that are cultivated for ornamental purposes outside of their native habitats and ecological and geographical ranges.

Results and Discussion

The results for the LC-MS profiling of saponins in extracts of *Silphium perfoliatum* from Poland and the United States are provided in Table 5.2. A direct comparison of these results is not possible because the data were collected using different HPLC and MS instruments and methods. Still, there is valuable information that can be obtained by comparing the two separate analyses. The extracts of *S. perfoliatum* grown in

[5] The chemical analyses of *S. perfoliatum* extracts from plants cultivated in Poland were produced through the joint efforts of Malgorzata Wojcinska (extraction and purification), Klaus Linse (LC-MS instrumentation), Schonna Manning (sample preparation and LC-MS instrumentation), and Lalita Calabria (LC-MS data analysis).

Poland are not identical to the extracts of *S. perfoliatum* collected from its native habitat in the United States. However, five of the compounds detected in both LC-MS analyses share the same order of elution, molecular weights and in some cases, the same fragment ions.

United States			Poland		
Cmpd #	**Rt**	**ESI-MS**	**Cmpd #**	**Rt**	**ESI-MS**
1	30.55	941.5 [M-H]-			
2	34.54	809.6 [M-H]-, 647.3 [M-hex-H]-, 471.5 [M-hex-gluA-H]-	**1**	13	809.7 [M-H]-, 647.4 [M-hex-H]-, 471.2 [M-hex-gluA-H]-
3	36.07	925.6 [M-H]-	**2**	13.3	809.6 [M-H]-, 633.4 [M-gluA-H]-, 471.3 [M-gluA-hex-H]-
4	36.63	955.6 [M-H]-, 793.6 [M-hex-H]-			
5	37.98	925.6 [M-H]-, 1851.8 [2M-H]-, 763.6 [M-hex-H]-			
6	40.99	793.6 [M-H]-, 1587.9 [2M-H]-, 631.7 [M-hex-H]-, 455.7 [M-hex-gluA-H]-	**3**	14.1	793.9 [M-H]-, 617.6 [M-gluA-H]-, 455.6 [M-gluA-hex-H]-
7	42.18	793.6 [M-H]-, 617.8 [M-gluA-H]-, 455.4 [M-gluA-hex-H]-	**4**	14.6	793.9 [M-H]-, 631.5 [M-hex-H]-, 455.5 [M-hex-gluA-H]-
8	43.11	835.5 [M-H]-, 673.5 [M-hex-H]-, 569.6	**5**	16.3	795.7 [M-H]-, 633.5 [M-hex-H]-, 471.4 [M-hex-hex-H]-
9	45.17	941.7 [M-H]-, 1883.8 [2M-H]-	**6**	16.7	941.7 [M-H]-, 779.9 [M-hex-H]-, 617.6 [M-hex-hex-H]-
10	46.04	649.5 [M-H]-, 487.4 [M-hex-H]-	**7**	17.2	793.9 [M-H]-, 617.9 [M-gluA-H]-, 455.3 [M-gluA-hex-H]-
11	46.78	955.6 [M-H]-, 835.6, 793.6 [M-hex-H]-, 631.6 [M-hex-hex-H]-	**8**	17.6	983.7 [M-H]-, 821.9 [M-hex-H]-, 659.6 [M-hex-hex-H]-, 455.4 [M-agly-H]-
12	47.74	835.5 [M-H]-, 673.5 [M-hex-H]-, 647.5, 569.7	**9**	18.4	835.8 [M-H]-, 659.4 [M-gluA-H]-, 471.3 {M-agly-H]-
13	54.07	763.6 [M-H]-, 1527.8 [2M-H]-	**10**	19.9	793.9 [M-H]-, 631.6 [M-hex-H]-, 455.5 [M-hex-gluA-H]-
14	59.02	631.6 [M-H]-, 455.6 [M-gluA-H]-	**11**	20.7	779.6 [M-H]-, 617.4 [M-hex-H]-, 455.5 [M-hex-hex-H]-

Table 5.2. HPLC/ESI-MS negative mode data showing intraspecific variation of saponins for *S. perfoliatum* from the United States (left) and from Poland (right). Compounds shared between both geographic isolates are shaded. LC-MS analyses were performed on different instruments with different methods. Thus, the data presented in this figure are correlative and without authentic standards, only tentative assignments can be made based on the LC-MS profiles shown here.

With the exception of compound **12** in the United States isolate and compound **9** in the Poland isolate (*m/z* 835), the molecular weight and fragmentation of all other shared compounds strongly correlate with previously isolated triterpene saponins from *S. perfoliatum* (silphioside E, *m/z* 941; silphioside G, *m/z* 793), *S. integrifolium* (cynarasaponin C, *m/z* 793; see section 5.3.1 of Chapter 5), and *S. morhii* (ilexoside XLVIII, *m/z* 809; see section 5.3.2 Chapter 5) (Davidyants et al., 1984 a, b, c). The disparities in the compound retention times between the Poland and Unites States LC-MS data reflect the differing gradients and column types used in each analysis and do not necessarily correspond to differences in a given compounds polarity. Although, tentative assignments of these compounds can be made, the authors understand that without authentic standards it is not possible to confirm the exact structures for these compounds.

The intraspecific variation observed between *S. perfoliatum* cultivated in Poland and *S. perfoliatum* collected from its native habitat in the United States is probably a result of geographical differences (latitude, longitude, length of day, etc.) and perhaps, the effects of cultivation vs. wild grown conditions (nutrients, exposure, moisture, etc). The intraspecific variation of some of the saponins in *S. perfoliatum* seems to be greatly dependent on environmental factors. These results are not surprising considering that many of these compounds serve a protection role against insects and pathogens, and in some cases, have profound physiological effects on growth and development. A separate LC-MS analysis of *S. perfoliatum* saponins cultivated in Poland carried out by Kowalski (2007) showed an even more divergent saponin profile compared with both samples tested in the current analysis. The compounds present in both geographical isolates and in

Kowalski's (2007) analysis are probably influenced more by genetic and/or developmental factors than compounds that are expressed only in one of the geographical isolates. Future studies are needed to fully understand the functional importance of these compounds in *Silphium* and for plants in general.

5.3.4 Experimental

General experimental procedure- ESI-MS was performed on a Finnigan LCQ Ion Trap mass spectrometer and HR-ESI-MS experiments were performed on Fourier Transform Ion Cyclotron mass spectrometer (Ion Spec, Varian). The spectra were recorded by infusion into the ESI (Electrospray Ionization) source using MeOH as a solvent. ^{1}H and ^{13}C NMR spectra were recorded in either pydridine-d_5 or CD_3OD on a Varian Inova 500 Mz spectrometer. All chemical shifts (δ) are given in ppm units with reference to TMS as an internal standard and the coupling constants (J) are given in Hz.

Column chromatography was carried out on Kieselgel 60 (60-200µm, Merck) or LiChroprep RP-18 (40-60µm, Merck). MPLC was performed on a Beckman 120B pump using Millipore Vantage-L (16 x 250, 32 x 250 mm) columns with a flow rate of 2ml min^{-1}. Fractions were monitored by TLC on silica gel plates (Merck precoated silica gel 60 F_{254}) and developed in the solvent system, EtOAc: HOAc:H_2O (9:2:2). Spots were visualized by heating after spraying with 10% H_2SO_4.

Plant material - *S. integrifolium* was collected on July 12, 2002 in Bowie Co., TX, approximately 10.5 miles southwest of the community of New Boston. The plant material

was identified by Dr. Jeffrey Williams and Dr. Malgorzata Wojcinska and vouchered by Dr. Mark Bierner. A specimen was deposited as No. *JW20021712* at the University of Texas at Austin. *S. morhii* leaves were collected on June 19, 2003 in Dade Co., GA along GA Interstate 136, 3.2 mi east of Trenton in an open area near the woods, growing with S. *asteriscus.* The plant material was identified by Dr. Jennifer Clevinger and a voucher specimen (No. 221) was deposited at The University of Texas at Austin. *S. perfoliatum* leaves were collected from the Garden of Medicinal Plants at the University of Medical Sciences in Poznan, Poland in June 2002. The leaves were dried at room temperature prior to extraction.

Extraction and isolation- The dried mixture of leaves and stems of *S. integrifolium* or *S. morhii* (~500 g) were successively extracted for 48 hours with CH_2Cl_2-MeOH (1:1), MeOH and MeOH-H_2O (1:1) at room temperature to give three separate crude residues. The conc. MeOH and MeOH-H_2O extracts were combined and dried under vacuum and the concentrate (70 g) was dissolved in water and loaded onto a 7.5 x10 cm C-18 column. The column was washed with water to remove sugars and then with 40% MeOH to remove phenolics. Saponins were removed with 85% MeOH and the crude saponin fractions were subjected to reversed phase C-18 MPLC using a MeOH:H_2O gradient to give 5 major fractions (I-V) for *S. integrifolium* and 4 major fractions for *S. morhii*. For *S. integrifolium*, fractions I-III were repeatedly subjected to reversed phase C-18 MPLC purification on silica RP-18 using a MeOH-H_2O gradient, affording compounds **1** (34 mg) and **2** (66 mg). Remaining saponin mixtures IV and V obtained from the above

processing were further subjected to silica gel CC eluted with a isocratic system of EtOAc: AcOH: H_2O (9:2:2) affording compounds **3** (52mg) and **4** (35mg). For *S. morhii*, fractions I-V were subjecting to reversed phase C-18 MPLC using a MeOH:H_2O gradient, affording several subfractions, which were further purified by silica gel CC eluted with a isocratic system of EtOAc: AcOH: H_2O (9:2:2) affording compound **1** (6mg) and several partially purified saponin mixtures.

Dried *Silphium perfoliatum* leaves (1g) were extracted with CH_2Cl_2-MeOH (1:1) at room temperature twice, followed by extraction with 50% MeOH (twice; also at room temperature). The extracts were filtered, and the residues combined. The solvent was removed under pressure. The combined crude extract was suspended in water and passed through a C-18 column preconditioned with water. The column was washed with water to remove sugars and then with 50% methanol to remove phenolics. Most of the saponins were removed with 80% methanol. However, some saponins eluted in water and with 50% MeOH. All fractions were analyzed by LC-MS.

HPLC/ESI-MS- S. perfoliatum leaf extracts were dissolved in analytical grade MeOH prior to injection. Tandem mass spectrometry was carried out using a micro-capillary gradient LC system (1100 series, Agilent Technologies, San Jose, CA) equipped with an Esquire-LC ion-trap (Bruker Instruments, Bellerica, MA). A reversed phase C-18 MS column (Vydac 218 MS 3.1505, 150 μm x 5 cm) was used, with the flow rate of 2 μl/min. The binary mobile phase consisted of solvent A (0.1% HCO_2H, 0.01% TFA in H_2O) and solvent B: 0.1% HCO_2H, 0.01% TFA in H_2O/CH_3CN (10:90). The column was

equilibrated at 5% B and run using a gradient of 5% B (0 min), increasing to 95% B in 56 min; a column clean-up stage was performed at 5% B (3 min). The effluent from the HPLC column, monitored by a UV detector at 254 nm, was directed on-line into the orthogonal Esquire-LC electrospray source. The ion spray experiment was performed in negative mode; endplate offset was – 500 V; ions were scanned from 50 to 1500 *m/z*; the nebulizer gas was set to 23.0 psi, the dry gas to 7.00 L/min; the drying temperature at the capillary entrance was 150 °C; capillary exit was set to 76.1 V; skimmer 1 at 26.1 V, trap drive to 58; an average of 3 spectra were acquired over a time period of 100 ms; the collision gas was helium; the MS/MS experiment was performed in the auto mode.

Cell Proliferation Assay - MDA-MB-231 human breast cancer cells were cultured as described in Calabria et al., 2008. 10,000 cells per well in a total volume of 100 μL of media with 5% FBS were cultured for 24 h in a 96 well plate (9 replicates/treatment) followed by treatment with vehicle (0.1% DMSO) or saponins at 25, 50, or 100 μg/ml in DMSO for 48 h. Cell proliferation was quantified using the CellTiter 96 AQ_{ueous} Non-Radioactive Cell proliferation Assay (Promega Corp., Madison, WI) according to manufacturer's instructions. This assay uses a tetrazolium compound that is bioreduced by metabolically active cells into a formazan product. The absorbance of the soluble formazan product in the tissue culture medium as measured at 490 nm is proportional to the number of viable cells.

Chapter 6: Metabolic Profiling of Triterpene Saponins in the Genus *Silphium* and the Closely Related Species *Lindheimera texana* by Electrospray Ion-Trap Mass Spectroscopy[6]

Triterpene saponins from aerial parts of all species in the genus *Silphium* L. (Asteraceae) and the closely related species *Lindheimera texana* were profiled and characterized using reverse phase HPLC with electrospray ionization mass spectrometry (ESI/MS) to examine phytochemical variation among the studied species and to examine the chemosystematic value of saponins in these taxa. Over 100 saponins comprising 13 previously reported and 89 newly detected compounds were tentatively identified based on negative ionization HPLC/ESI-MS analysis and comparisons with authentic standards previously reported for *S. radula* as well as other literature data. In addition, mass spectral characteristics of nine standards from *S. radula* were also analyzed by HPLC/ESI/MS/MS. The metabolic profiling of triterpene saponins from all *Silphium* species provides the first comprehensive profiling of saponins in this genus.

6.1 Introduction

Silphium L. (Asteraceae) is a small genus of eleven species common to the prairies and woodlands of central North America (Clevinger and Panero, 2000). Native American tribes, including the Cherokee Indians, considered *Silphium* a medicinal herb and prepared extracts of leaves and flowers for respiratory and kidney ailments; they also

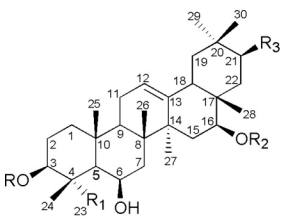
[6] The LC-MS analyses described in this chapter were carried out in the laboratory of Dr. Paul W. Paré at Texas Tech University in Lubbock, Texas. The data described here was produced by the joint efforts of Lalita M. Calabria, Venkat Krishnamachari, Paul W. Paré and Tom J. Mabry and is prepared in a form to be submitted for publication as: Metabolic profiling of triterpene saponins in *Silphium* and *Lindheimera texana* by electrospray ion-trap mass spectroscopy.

used the roots of *Silphium* as a poultice for bleeding wounds, backaches and hemorrhaging (Hamel and Chiltoskey, 1975). Furthermore, biological screening of North American prairie plants with ethnobotanical uses concluded that organic extracts of leaves, stems and roots of *S. perfoliatum* and *S. laciniatum* exhibited moderate activity in anti-cancer screening (LC_{50} > or = 20% of the all cell lines tested); aqueous extracts of *S. laciniatum* showed potent anti-HIV activity (LC_{50} > 50% compared with controls); moreover, organic extracts of *S. perfoliatum* showed moderate activity (LC_{50} <50% of control) in anti-HIV assays (Kindscher et al., 1998).

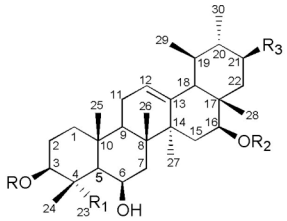
Phytochemical screenings of extracts from a few *Silphium* species resulted in the isolation of almost all traditional classes of secondary metabolites, including flavonoids (El-Sayed et al., 2002), phenolic acids (Kowalski and Wolski, 2003), essential oils (Kowalski et al., 2005), sesquiterpenes (Bohlmann and Jakupovic, 1979; 1980), diterpenes (Pcolinski et al., 1994), and an abundance of oleanolic-type triterpene saponins (Davidyants et al., 1984a,b; Davidyants et al., 1985; Davidyants et al., 1986; Calabria et al., 2008).

Saponins are an economically important group of natural products with a wide range of biological activities and a broad distribution in the plant kingdom. Typically, saponins consist of one or more sugar moieties attached to either a steroid or a triterpene aglycone. Structures of saponins reported from various *Silphium* species are provided in Figure 6.1 and Figure 6.2. Several studies have been published describing a range of biological activities for these saponins. For example, the cytotoxic activities of polyhydroxylated pentacyclic oleanene and ursene-type triterpene saponins from leaf

methanolic extracts of *Silphium radula* were tested against the human breast cancer cell line 25 MDA-MB-231 and showed that saponin **I**, decreased cell proliferation in a statistically significant manner at 25 μg/ml (Calabria et al., 2008).



I



II

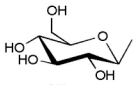
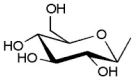
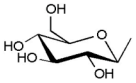
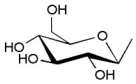
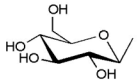
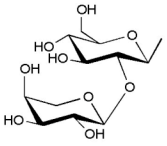
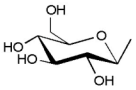
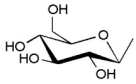
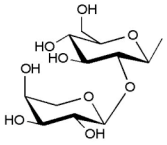
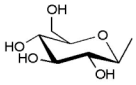
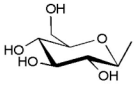
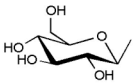
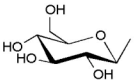
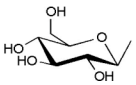
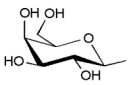
MW	Saponin	Aglycone	R	R_1	R_2	R_3
814	A	I		CH_2OH		OH
650	B	I		CHO	H	OH
636	C	II		CH_3	H	OH
930	D	I		CH_2OH		H
636	E	I		CH_3	H	OH
914	F	I		CH_3		H
812	G	I		COOH		H
796	H	I		CHO		H
782	I	II		CH_3		H

Figure 6.1. Structures of triterpene saponins isolated from *S. radula* (Calabria et al., 2008). Trivial letter designations (A-I) for each saponin are used for reference throughout this chapter.

A mixture of saponins isolated from *S. perfoliatum* (saponins **J-O**, Figure 6.2) exhibited blood cholesterol-lowering activity when administered orally in both normal and hyperlipidemic rats (Syrov et al., 1992) and significantly inhibited the growth of the phytopathogenic fungi, *Dhreslera graminea* as well as inhibiting both mycelial growth and spore formation of *Rhizopus nodosus* and *R. nigricens*. (Davidyants et al., 1997). More recently, the growth-regulating effects of saponins from *S. perfoliatum* were reported; a saponin mixture showed cytokinin and auxin-like activities on seed germination, hypocotyl elongation and root formation (Davidyants et al., 2006).

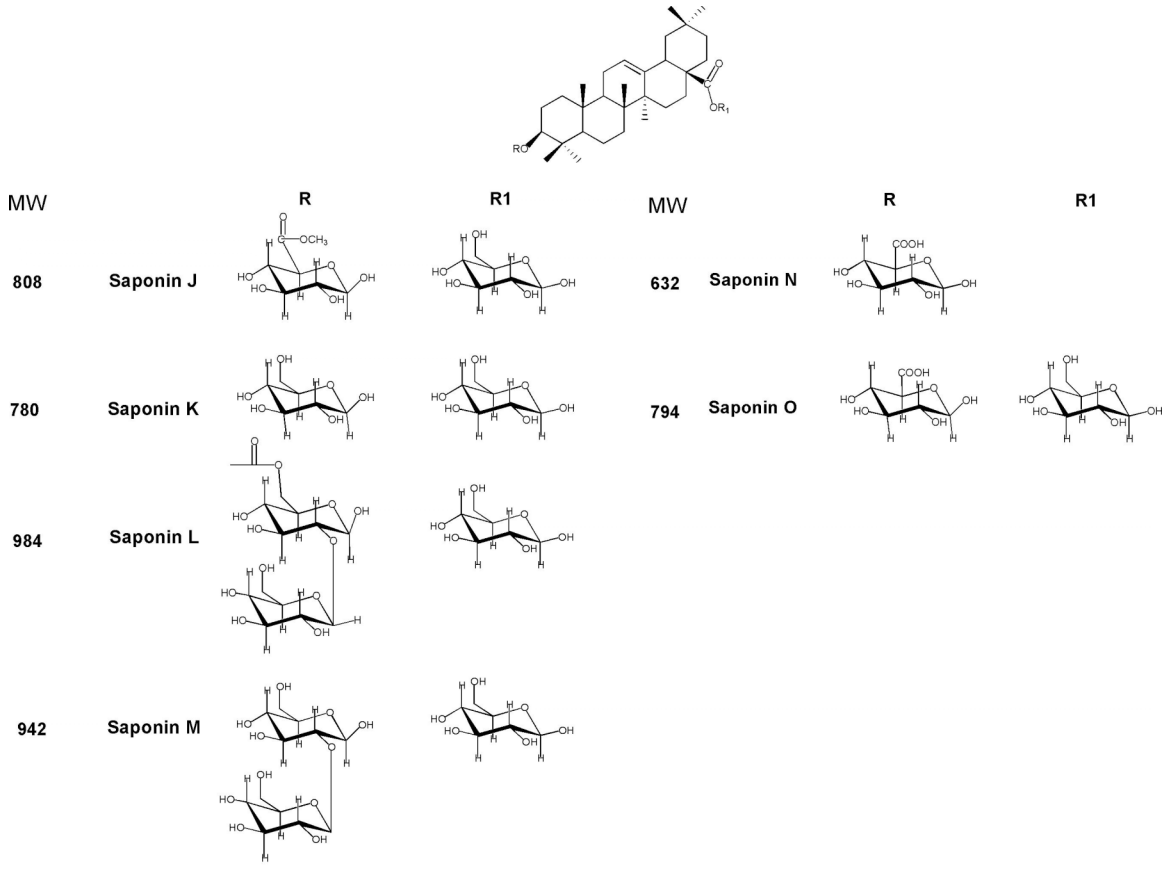


Figure 6.2. Structures of triterpene saponins isolated from *Silphium perfoliatum* (Davidyants et al., 1984; Davidyants et al., 1984b; Davidyants et al., 1985; Davidyants et al., 1986). Trivial letter designations (J-M) assigned for each saponin are used for reference throughout this chapter.

Saponins possess both hydrophobic and hydrophilic properties and often occur as complex mixtures in plants, making their isolation both time-consuming and complicated. HPLC has become a popular method for overcoming some of the difficulties encountered with the determination of saponin mixtures and many HPLC techniques have been reported in the literature (Oleszek, 2002). Unfortunately, difficulties with detection and solvent restrictions frequently outweigh the speed and sensitivity often associated with HPLC methods because most saponins lack a chromophore suitable for UV detection. Derivatization of saponins with an appropriate chromophore can facilitate UV detection at higher wavelengths, but the absorption spectra of derivatized saponins are still not specific enough to ensure full identification. However, the combination of HPLC with mass spectrometry (MS) detection has proven to be a favorable method for structural analysis and characterization of saponins. In particular, several studies have indicated electrospray ionization mass spectrometry (ESI) coupled with HPLC is well suited for the identification of saponin mixtures in plant extracts (Perret et al., 1999; Huhman and Sumner, 2002; Guo et al., 2002; Kapusta et al., 2005; Madl et al., 2006; Li et al., 2007). In ESI-MS, characteristic fragmentation patterns are observed and often allow the determination of aglycones, saccharide sequences, and with ESI-MS^n, additional aglycone fragments can provide additional structural information. Huhman and Sumner (2002) provided an excellent example of this method in their report on the metabolic profiling of saponins from *Medicago sativa* and *Medicago truncatula* using HPLC/PDA/ESI/MS. Their analyses utilized both UV and MS detection and additional tandem mass detection (MS^2) to aid in structural determination, resulting in the

identification of 15 saponins in *M. sativa* and 27 in *M. truncatula*. However, only a limited number of standards from *M. sativa* were available, making full characterization of all saponins not possible. Kapusta et al. (2005) reported a similar HPLC/PDA/ESI/MS analyses of three subspecies of *M. truncatula*, but in contrast to Huhman and Sumner (2002), their method was based on 18 standards previously isolated from *M. truncatula*, which allowed for much higher level of confidence when discriminating between structurally different compounds having identical masses.

Only one study has been published on the analysis of saponins from *Silphium* using LC-MS (Kowalski, 2007). Kowalski's study focused on the qualitative and quantitative evaluation of triterpene aglycones from saponin fractions of three species of *Silphium* (*S. perfoliatum, S. trifoliatum* and *S. integrifolium*) as compared with *Panax quinquefolium* root and *Calendula officinalis* flowers using HPLC-PDA/ESI/MS. The results from this study revealed the presence of 36 oleanolic and ursolic acid saponins in leaves, stems and roots of the *Silphium* species examined. However, Kowalski's (2007) analysis did not include authentic standards from *Silphium* species, and therefore, a full characterization of the saponins was not obtained.

Here we report the first comprehensive HPLC/ESI-MS analysis of saponins in all *Silphium* species. Additionally, *Lindheimera texana* was also analyzed as a species outgroup to better understand saponin distribution at the subtribal level. The present study presents a rapid and complete method for the analysis of saponin variation and distribution in the *Silphium* genus. Moreover, the chemosystematic value of saponins, a chemical group which is not well delineated at the genera or familial level and for which

there is very little known about their evolutionary roles, is examined in this genus as a starting point for future studies on the triterpene chemistry of the Asteraceae.

6.2 Materials and Methods

Plant material- Aerial parts of eleven *Silphium* species were collected from their native habitats across the United States by Dr. Jennifer Clevinger, Dr. Jeffery Williams and Lalita Calabria (PhD. candidate). The plant material was verified and voucher specimens for each species were deposited at the University of Texas at Austin (voucher numbers are in parentheses after species names). Aerial parts of *S. radula* (No. JW2002/7/3) were collected in Lee County, Texas in July 2002 near a rest area west of Giddings, Texas. *S. albiflorum* (No. JW2002/7/1) aerial parts were collected in Bosque C., TX on a dry upland Juniper savannah in July 2002. Aerial parts of *S. laciniatum* (No. 232) and *S. asteriscus* (No. 219) were collected near a roadside clearing in Greene Co., AL south of the town of Eutaw. *S. compositum* aerial parts (No. 228) were collected in the Cherokee National Forest, Polk Co., TN in June 2003. *S. terebinthinaceum* (No. 218) aerial parts were collected in a pine-oak forest near the town of Trenton, Dade. Co. GA in June 2003. Aerial parts of *S. brachiatum* (No. 224) were collected in the deep shade along the edge of the woods near the town of Sewance in Franklin Co., TN during June 2003. *S. integrifolium* aerial parts (No. JW2002/7/2) were collected in Hunt Co., TX at the Mathews Prairie Preserve in July 2002. Aerial parts of *S. morhii* (No. 221) were collected in Dade Co., GA from an open area near the woods in June 2003. Aerial parts of *S. perfoliatum* (No. 229) were collected in Giles Co., VA. along the New River in June

2003. *S wasiotense* aerial parts (No. 217) were collected in Norris Co., TN in Norris Dam State Park along the edge of the woods in June 2003. Aerial parts of *Lindheimera texana* (No. 003-Austin, personal collection) were harvested in the author's backyard in Austin, Texas during June 2006. All plant collections were air-dried, finely powdered and used for extraction.

Sample preparation- Dried and finely powdered *Silphium* or *Lindheimera texana* leaves and stems (1g) were extracted with 80% MeOH (50ml) at room temperature, overnight. The extracts were filtered and the same plant materials were extracted twice by refluxing with 80% MeOH (50ml) for 1 hr. Extracts were combined, and the solvent removed *in vacuo*.

Purification- The crude extract of each species was suspended in water and passed through a C-18 column preconditioned with water. The column was washed with water to remove sugars and then with 40% methanol to remove phenolics. Saponins were removed with 80% methanol and the solution was concentrated *in vacuo* to remove solvent before weighing. All *Silphium* and *Lindheimera* extracts were resuspended in methanol to a final concentration of 10μg/μl.

High performance liquid chromatography- HPLC analyses were performed on a Thermo Separation Products HPLC system coupled with a photodiode array (PDA) detector and a Thermo Finnigan (San Jose, CA) LCQ quadrupole mass spectrometer in sequence. An

Alltech Altima C18 column (250mm × 2.1mm, 5μm) was used for separation using a mobile phase gradient consisting of water (A) and acetonitrile (B). Acetic acid (0.1%) was used as a mobile phase modifier to facilitate ionization and to achieve a high level of chromatographic separation efficiency and ESI-MS instrument response. The system was run with the following gradient program: from 5-95%B in 56 minutes, from 95-100%B in 2 minutes and held for 4 minutes before returning to 5%B in 3 minutes. A flow was set at 0.2 ml/min and directed to a quadrapole ion trap mass spectrometer via an ESI interface. The mass spectrometer was operated in negative mode. MS-MS was performed for all *S. radula* standards. A sample injection volume of 10μl was used for all analyses performed.

Mass Spectrometry- The negative ion ESI-MS experiments were performed using a Thermo Finnigan (San Jose, CA) LCQ quadrupole mass spectrometer equipped with an electrospray ion source. The ESI conditions were as follows: spray voltage, 4.5 kV; capillary offset voltage -45 V; capillary temperature 200.00 °C. Mass spectra were recorded in the range of 400 - 2000 *m/z*.

Triterpene saponin standards- The nine saponin standards used in this study (compounds **A-I**) were previously isolated from *Silphium radula* and their structures were determined by NMR and HR-ESI-MS and other chemical analyses (Calabria et al., 2008). A standard mixture containing 1μg/μl of **A-I** was used for calibration of the mass spectrometer.

6.3 Results and Discussion

An HPLC/ESI-MS method was established for the profiling of saponins from all *Silphium* species as well as *Lindheimera texana* to allow for the qualitative differentiation and characterization of these compounds and to examine the chemosystematic value of saponins in these taxa at the species and genera level. Previous studies describing the isolation and identification of saponins from *Silphium radula* provided the necessary standards to optimize the HPLC/ESI-MS profiling method (Calabria et al., 2008). In addition, previously reported structural data and LC-MS analysis for saponins isolated from *S. perfoliatum*, *S. integrifolium* and *S. asteriscus* provided additional information for the preliminary identification of *ca*. 90 unreported metabolites (Kowalski, 2007).

6.3.1 Profiling of *S. radula*

HPLC/ESI-MS analyses of *S. radula* standards and extracts served as positive controls to establish HPLC retention times and characterize the mass spectra of known *Silphium* saponins. HPLC retention times and observed ESI/MS and ESI/MS/MS fragmentation patterns of saponins **A-I** are provided (Table 6.1).

Retention times and fragmentation patterns from HPLC/ESI-MS analyses of *S. radula* saponins were used to confirm the identity of saponins for the remaining *Silphium* species. Based on previous mass spectral studies on saponins from *S. radula* (Calabria et al., 2008), negative ion mode was chosen for the HPLC/ESI-MS profiling of *Silphium* species due to the greater sensitivity and higher signal-to-noise ratio than was observed in positive mode.

Saponin	Rt	Cmpd #	Aglyc	MW	MS	MS/MS
A	32.27	2	490	814	813.7 (100) [M-H]-, 1627.8 (8) [2M-H], 651.5 (12) [M-Glc-H]-, 489.3 (3) [M-Glc-Glc-H]-, 471 (5) [Agly-H20-H]-	651, 633, 585, 489, 471, 454
B	35.61	22	488	650	649.4 (52) [M-H]-, 709.4 (100) [M-H+59]-, 1299.6 (33) [2M-H]-, 487.4 (9) [M-Glc-H]-, 559.5 (33) [M-90-H]-	487, 469, 437, 421
C	36.61	28	474	636	635.5 (34) [M-H]-, 1271.7 (47) [2M-H]-, 695.4 (100) [M-H+59]-, 473.5 (1) [M-Glc-H]-	616, 588, 472, 454
D	37.8	30	474	930	929.6 (100) [M-H]-, 1859 (5) [2M-H]-, 797.5 (7) [M-Ara-H]-, 779.7 (2) [M-Ara-H20-H]-, 635.6 (1) [M-Ara-Glc-H]-, 473 (0.4) [M-Ara-Glc-Glc-H]-	797, 779, 635, 617, 569, 473
E	38.18	32	474	636	635.6 (28) [M-H]-, 695.4 (100) [M-H+59]-,1271.6 (45) [2M-H]-, 473.3 (1) [M-Glc-H]-	616, 588, 553, 363
F	38.91	34	458	914	913.7 (100) [M-H]-, 1827.8 (13) [2M-H]-, 781.7 (3) [M-Ara-H]-, 619.6 (1) [M-Ara-Glc-H]-, 457.3 (1) [M-Ara-Glc-Glc-H]-	781, 763, 619, 601, 423
G	43.1	54	488	812	811.6 (100) [M-H]-, 1623.8 (32) [2M-H]-, 649.8 (1)[M-Glc-H]-	751, 649, 589
H	43.46	57	472	796	795.7 (100) [M-H]-, 1591.8 (9) [2M-H]-, 633.4 (3) [M-Glc-H]-	733, 633, 571, 469
I	53.53	92	458	782	781.7 (100) [M-H]-, 1564.6 (20) [2M-H]-, 617.1 (1) [M-Glc-H]-, 455.3 (1) [M-Glc-Gal-H2O-H]-	619, 601, 453

Table 6.1. LC-MS retention times and mass spectral characteristics for saponins A-I, all previously reported from *Silphium radula* (Calabria et al., 2008).

Negative-ion mass spectra yielded strong $[M-H]^-$ and $[2M-H]^-$ ions for *S. radula* saponins that allowed for easy confirmation of molecular weights. Moreover, by maintaining a low capillary voltage offset, negative fragment ions corresponding to the sequential loss of sugars were observed, giving additional structural information to aid in the identification of unknown compounds. In addition, HPLC/MS/MS was used to detect further fragmentation of the standards. We did not perform MS^2 studies on *Silphium* or *Lindheimera* extracts because of the large amount of data acquisition required for the analysis of all twelve species containing potentially hundreds of compounds.

The chromatographic behavior of the polyhydroxylated saponins isolated from *S. radula* was assessed. The order of retention times was directly correlated with the number of free hydroxyls groups and number of glycosides attached to the aglycone. For example, the earliest eluting saponin (*m/z* 813) had three hydroxyl groups present (C-6, C-22 and C23), and two sugars attached at the C-3 and C-16 position, resulting in increased Van der Waal interactions, where the hydrophilic carbohydrates and hydroxyl groups shield the hydrophobic aglycone from the hydrophobic stationary phase. This phenomenon is further demonstrated by the last eluting saponin (*m/z* 781), which only has one free hydroxyl (C-6) and two sugars attached at the C-3 and C-16 position. Figure 6.3 provides examples of three mass spectra for standards isolated from *S. radula.*

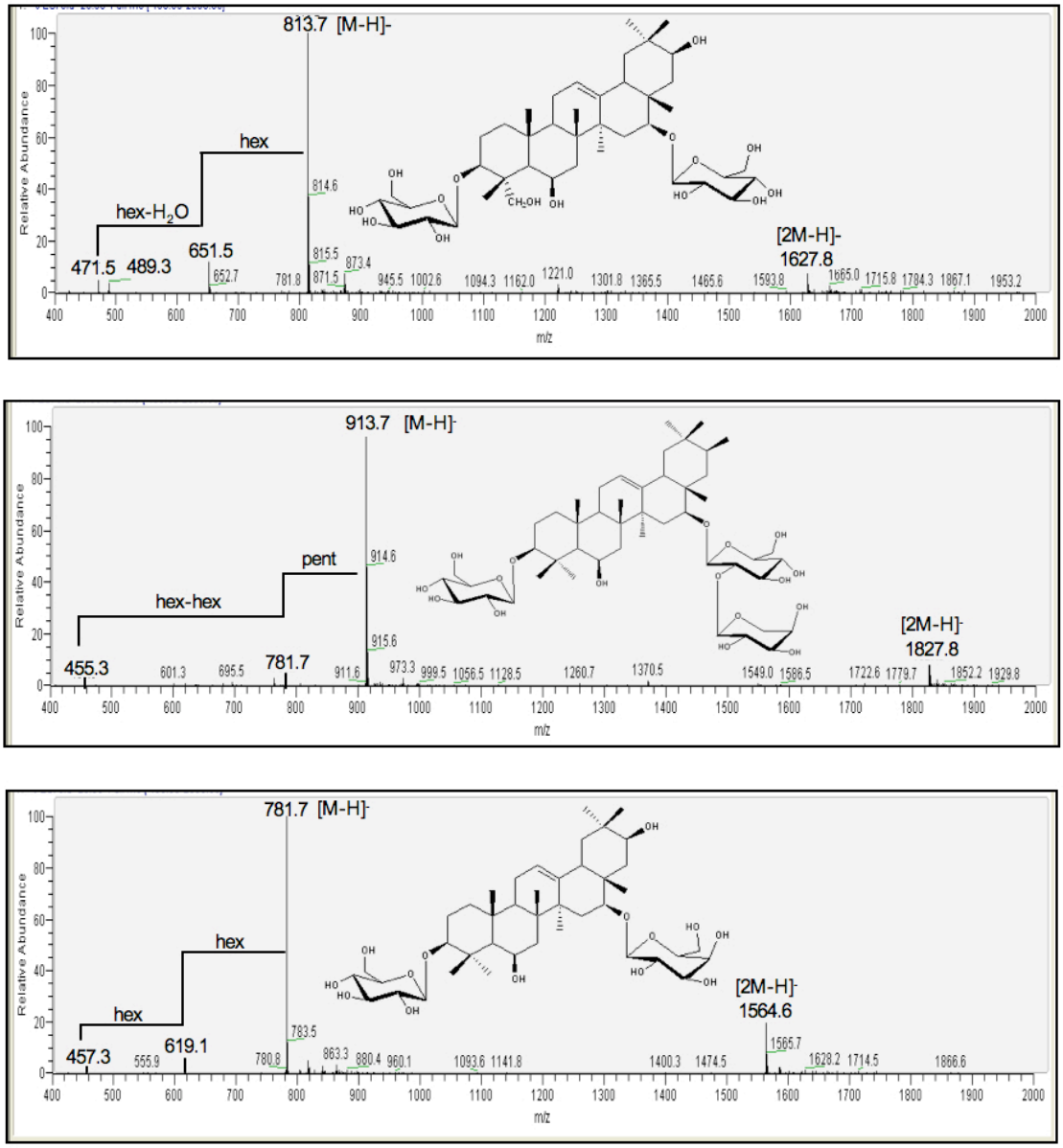


Figure 6.3. MS spectra of the [M-H]$^-$ of literature reported saponins, A (*m/z* 813), F (*m/z* 913) and I (*m/z* 781) from *Silphium radula* (Calabria et al., 2008). [M-H]$^-$ was the most abundant ion for saponins A, F and I. Additional ion fragments were observed corresponding to the loss of sugar moieties as well as dimers.

6.3.2 Profiling Saponins from all *Silphium* Species and *Lindheimera texana*

The optimized, negative-ion HPLC/ESI-MS parameters were used to profile and compare saponins in extracts from *Silphium* species and *Lindheimera texana*. The results of our HPLC/ESI-MS analysis of triterpene saponins from all *Silphium* species showed considerable differences in the number of saponins from species to species. Base peak chromatograms for all species analyzed in this study are shown in Figures 6.4a, 6.4b and 6.4c. The peaks observed between Rt=20-24 minutes in some chromatographic traces correspond to polar flavonoids which were not fully removed in the SPE clean-up step of the crude *Silphium* and *Lindheimera* extracts. Although most flavonoids eluted between 40-60% methanol, a few polar flavonoids remained in the 85% methanol eluate (saponin-containing fraction). However, this did not interfere with our LC-MS analysis because the earlier retention times and differing spectral properties of flavonoids allowed for easy differentiation of peaks corresponding to saponins.

The great variation of saponins between species translated to a large variance in retention times for individual compounds. For example, for saponin **A**, the corresponding peak was detected between 30.05 an 33.06 minutes showing a variance of 2.99 minutes between four species. In addition, different species had clusters of compounds concentrated in different retention time windows, making it difficult to optimize a gradient that fully resolved all peaks in all species. But even though some compounds had overlapping retention times in the chromatographic traces, the saponins could be differentiated using extracted ion chromatograms and mass spectral fragmentation data.

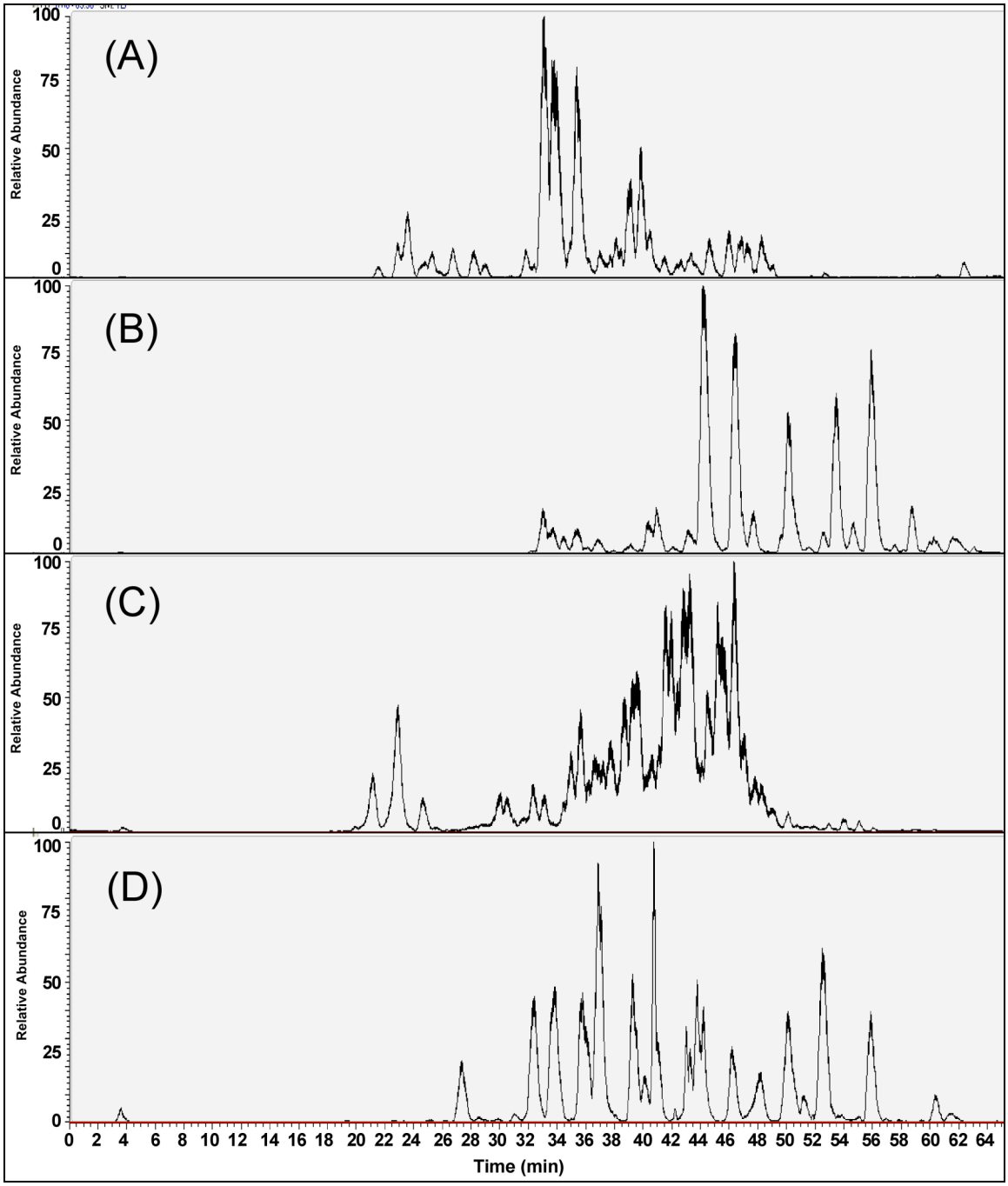


Figure 6.4a. Base peak chromatograms obtained by negative ion HPLC/ESI-MS of leaf extracts of (A) *Silphium asteriscus,* (B) *S. brachiatum,* (C) *S. integrifolium* and (D*) S. morhii.* Separation was achieved using injections of 10µg total extract, reverse-phase HPLC and gradient elution with a 0.1% aqueous acetic acid and acetonitrile. A complete list of saponins, including retention times is presented in Table C.1 (Appendix C).

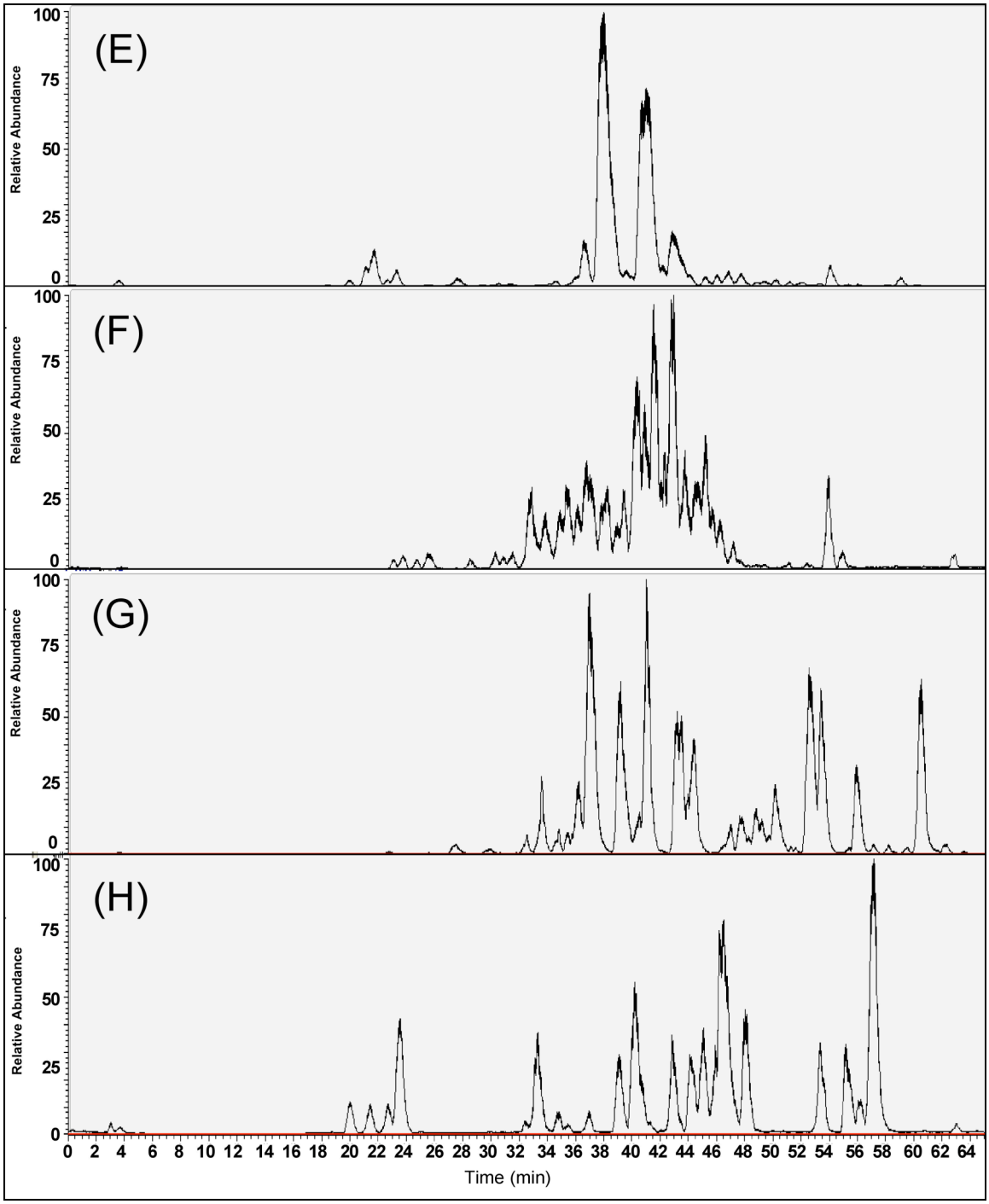


Figure 6.4b. Base peak chromatograms obtained by negative ion HPLC/ESI-MS of leaf extracts of (E) *Silphium perfoliatum*, (F) *S. radula*, (G) *S. wasiotense* and (H) *S. albiflorum*. Separation was achieved using injections of 10µg total extract, reverse-phase HPLC and gradient elution with a 0.1% aqueous acetic acid and acetonitrile. A complete list of saponins, including retention times is presented in Table C.1 (Appendix C).

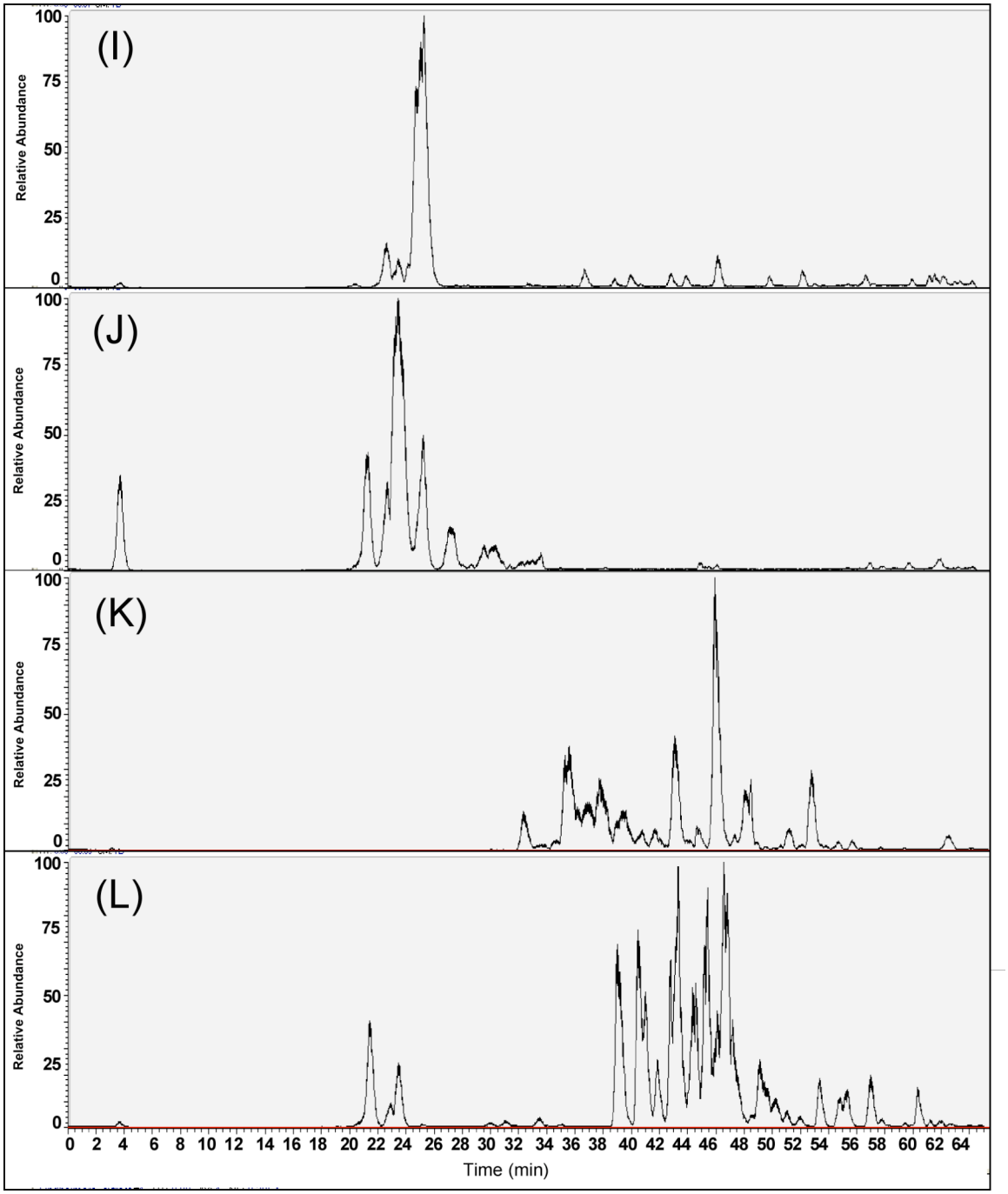


Figure 6.4c. Base peak chromatograms obtained by negative ion HPLC/ESI-MS of leaf extracts of (I) *Silphium laciniatum,* (J) *S. terebinthinaceum,* (K) *S. compositum* and (L) *Lindheimera texana.* Separation was achieved using injections of 10µg total extract, reversed-phase HPLC and gradient elution with a 0.1% aqueous acetic acid and acetonitrile. A complete list of saponins, including retention times is presented in Table C.1 (Appendix C).

Total saponin concentration (mg/g dry weight) in crude saponin extracts of aerial parts of *Silphium* species and *Lindheimera texana* are illustrated in Figure 6.5. Highest saponin concentration was observed in *S. radula* (130mg/g), with slightly lower concentrations in *S. integrifolium* (116.6mg/g), *S asteriscus* (105mg/g) *and S. compositum* (102mg/g). Between 9-27 saponins (mean=17.91) were detected in extracts of the *Silphium* species and *Lindheimera texana*, with the exception of *S. terebinthinaceum*, which showed only 3 saponins; the majority of this latter extract was composed of a ion with the molecular wt. of *m/z* 477 that fits well with the fragmentation pattern and retention time of an isorhamnetin monoside (Williams, 2006).

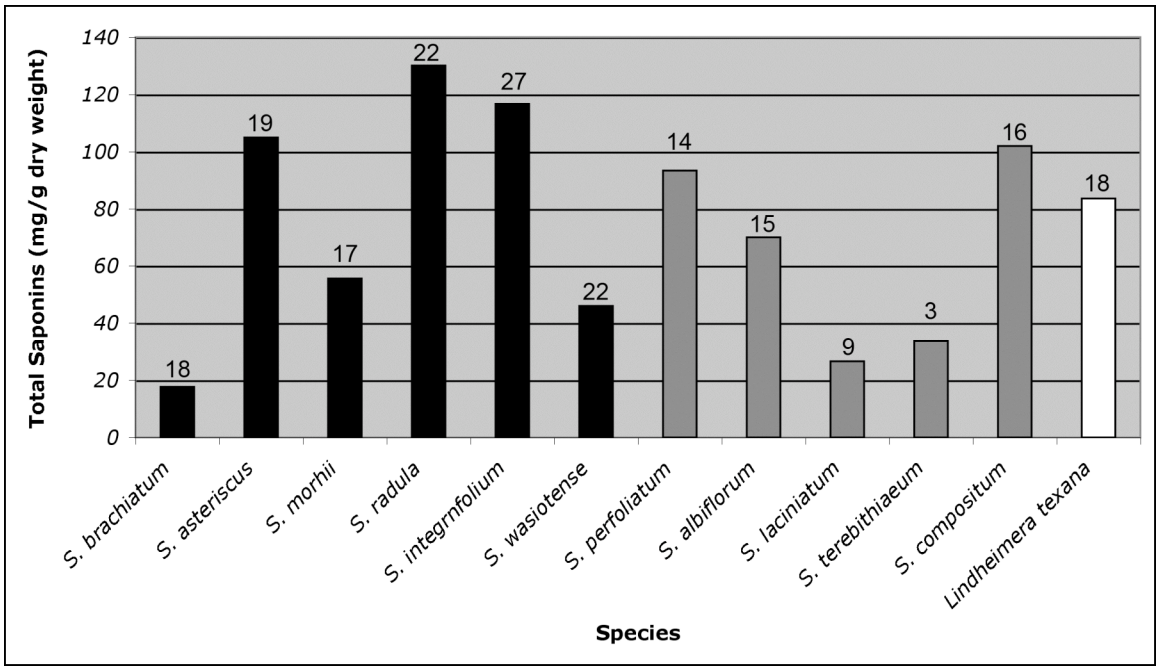


Figure 6.5. Total saponin concentration in crude extracts of *Silphium* and *Lindheimera texana.* Columns in black indicate species belonging to section *Silphium*, gray indicates section *Compositum* and white indicates the species outgroup, *Lindheimera.* Numbers above bars correspond to the total number of saponins detected for each species.

Species name	# of saponins per extract	3 most abundant saponins based on peak area, cmpd# (*m/z*)
S. radula	22	41 (959), 46 (943), 50 (985)
S. perfoliatum	14	24 (925), 47 (793), 55 (835)
S. albiflorum	15	24 (925), 73 (779), 100 (763)
S. asteriscus	19	2 (813), 10 (811), 18 (797)
S. brachiatum	18	60 (809), 73 (779), 97 (763)
S. integrifolium	27	48 (779), 49 (779), 72 (879)
S. morhii	17	13 (809), 58 (663), 90 (647)
S. wasiontense	22	25 (809), 45 (647), 90 (647)
S. compositum	16	53 (781), 68 (823)
S. terebinthinaceum	3	N/A
S. laciniatum	9	N/A
Lindheimera texana	18	44 (731), 47 (793), 74 (821)

Table 6.2. Summary table of HPLC/ESI-MS results for *Silphium* species analyzed in this study. A detailed list of all compounds (1-102), including retention times, observed ESI/MS ions and species is provided in Table C.1 located in Appendix C. Saponin abundance is based on the peak area calculated from base peak chromatograms of the corresponding $[M-H]^-$ ions. Peak area of saponin standards from *S. radula* were tested to confirm that the ionization responses for different saponins were relatively similar, that is, within the same order of magnitude.

Because of the relatively low concentrations of saponins detected in *S. terebinthinaceum* and S. *laciniatum* extracts when compared with all other species, a decision was made to examine their unusually large taproots, which are typically involved in the storage of carbohydrates and other metabolites (see Chapter 5.2 for results).

A summary table of results from the HPLC/ESI-MS profiling analysis for all tested saponin extracts is provided in Table 6.2. Saponin abundance in Table 6.2 is based on the peak area calculated from base peak chromatograms of the corresponding $[M-H]^-$ ions. Peak area of saponin standards from *S. radula* were tested to confirm that the ionization response for different saponins was relatively similar, within the same order of magnitude.

A comprehensive list of all compounds detected in *Silphium* species and *Lindheimera texana* is provided in Table C.1 (Appendix C), with retention times and corresponding mass fragmentation ions. One hundred and two saponins were detected in the analysis and 39 of these were shared between at least 2 species. Only saponin N (compound **102**), was present in all analyzed species, indicating this compound as a potential chemosystematic marker for *Silphium* and potentially other Asteraceae (Figure 6.6). *S. integrifolium* had the greatest number of saponins (27) and both *S. wasiotense* and *S. radula* also contained a large number of saponins (22) .

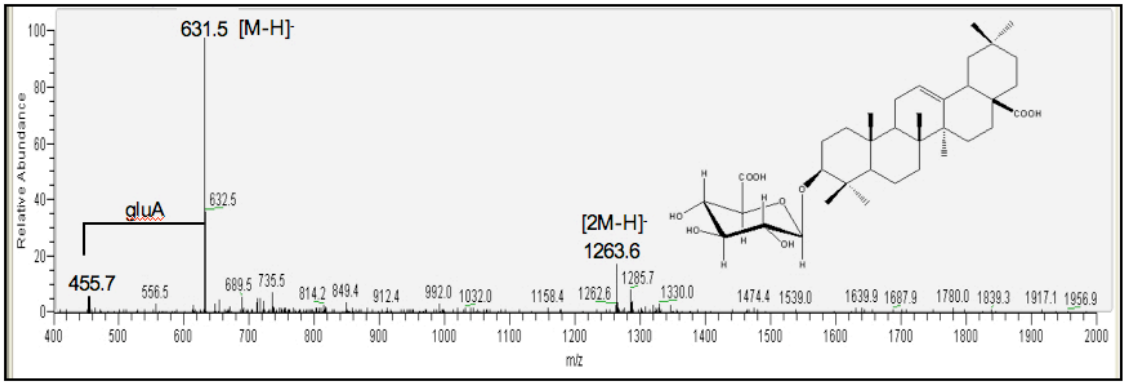


Figure 6.6. Mass spectrum for compound 102, tentatively assigned as saponin N, which was present in all species analyzed in this study.

Except for the fully characterized standards from *S. radula*, the tentative identifications of saponins in *Silphium* extracts were based on HPLC/ESI/MS data, co-migration times, comparison with mass fragmentation observed for the *S. radula* standards and literature references (Kowalski, 2007; Davidyants et al., 1984a, b,; Davidyants et al., 1985; Davidyants et al., 1986; Vidal-Ollivier et al., 1989; Li et al., 2007). Without authentic standards for these compounds, it is not possible to verify the

exact configurations or sugar attachments and therefore, the assignments made here are tentative and require further validation through classical isolation procedures and NMR spectral characterization.

There were some general trends observed when considering the Section level classification of *Silphium,* that is, section *Compositum* vs. section *Silphium*. There were significantly more saponins and a greater diversity of saponins detected in section *Silphium*, than found in section *Compositum.* However, this may be accounted for by the fact that section *Silphium* has 7 species and section *Compositum* has only 4. In addition, *S. laciniatum* and *S. terebinthinaceum*, the two species with the lowest quantity and number of saponins, are part of section *Compositum*. Even when considering these factors, almost every species from section *Silphium* produce a greater number of saponins overall. When comparing the LC-MS profiles of saponins between *Lindheimera texana* and *Silphium* species, ten of the eighteen saponins detected were present in at least one species of *Silphium*, which supports *Lindheimera's* close taxonomic relationship to *Silphium*.

S. radula- 22 peaks between the retention times of 32 and 55 minutes were detected in our LC-ESI/MS analysis of *S. radula*. All nine of the saponin standards previously isolated from *S. radula* were detected in the current analysis of *S. radula*. Of these standards, saponin **I** (compound **92**) was the most abundant saponin in the *S. radula* extract. Compound **102** could be tentatively assigned to saponin **N** previously reported from *S. perfoliatum*, based on the parent ion (*m/z* 631) and fragmentation patterns

observed in the mass spectra for this compound $[M-176-H]^-$, corresponding to the loss of a single glucuronic acid moiety. Compound **56** and **64** displayed the same molecular weight of *m/z* 779 and additional fragments that corresponded to the loss of two hexose units (*m/z* 617 and 455), which are in good agreement with the structure of saponin **K**.

S. perfoliatum- 14 peaks between the retention times of 30 and 59 minutes were detected in the *S. perfoliatum* extract. The majority of the saponins found in *S. perfoliatum* eluted in the retention time window of 36-43 minutes. None of the *S. radula* standards were detected in extracts of *S. perfoliatum*. The parent ion (*m/z* 793) and mass fragmentation of compound **47** are in good agreement with compound **O,** which displayed fragment ions corresponding to the loss of one hexose (*m/z* 631) and one glucuronic acid moiety (*m/z* 455). In addition, another molecular species present in our analysis of *S. perfoliatum,* compound **57,** displayed a parent ion (*m/z* 779) and fragmentation pattern (*m/z* 617, 455) in agreement with the structure of compound **K** previously reported from *S. perfoliatum*. In our analysis, the parent ion for compound **66** (*m/z* 941) showed a fragment ion (*m/z* 779) corresponding to the lost of a single hexose unit $[M-162-H]^-$. Further fragmentation was not present, however, based on previous findings from *S. perfoliatum*, we tentatively assigned compound **66** to saponin **M**. Finally, saponin **N** (**102**) previously reported from *S. perfoliatum* was also tentatively identified in our analysis of the *S. perfoliatum* extract.

It is interesting to note that saponins **J** and **L** previously reported from *S. perfoliatum* were not present in any *Silphium* extracts analyzed in this study. **J** and **L** are the only two saponins reported that possess an acetylated (for **L**) or methylated (for **J**)

sugar residue. These types of modifications can occur naturally, however, they are sometimes artifacts of harsh isolation procedures, such as boiling plant extracts in methanol. This method was employed during the extraction procedure reported in the original isolation of saponins **L** and **J** (Davidyants et al., 1984a). The purification of saponins is very difficult and can require several chromatographic steps, leading to unwanted modifications of the naturally occurring structures. Furthermore, in their original report (Davidyants et al., 1984a) for these compounds the authors did not include mass spectral data for the reported saponins. Therefore, the authors propose that a re-isolation and full chemical characterization of these compounds is warranted to establish whether saponins **J** and **L** are artifacts of compounds **M** and **O**, or whether they are naturally occurring derivatives.

S. albiflorum- 15 peaks between the retention times of 32 and 58 minutes were identified in our LC-MS analysis of the *S. albiflorum* extract. Saponins **A-I**, previously isolated from *S. radula*, were not detected in the *S. albiflorum* extract. Compounds **4** (*m/z* 941), **73** (*m/z* 779), **93** (*m/z* 793) and **102** (*m/z* 631**)** displayed identical molecular masses to saponins **M, K, O** and **N** previously isolated from *S. perfoliatum*. However, the mass fragmentation patterns and sugar types for the former three saponins differed from that of saponins **M, K, O**. Compound **102** showed the loss of a single glucuronic acid moiety [M-176-H]- and could be tentatively assigned to saponin **N**.

S. asteriscus- 19 peaks between retention times 33 and 60 minutes were detected in the *S. asteriscus* extract. Of the *S. radula* standards, compound **A** and **F** were also found in the *S. asteriscus* extract. Compound **47** (*m/z* 793), **56** (*m/z* 779), **64** (*m/z* 779) and **102** (631 *m/z*) matched the molecular masses and fragmentation patterns of saponins **O, K**, and compound **N**, previously reported from *S. perfoliatum*. However, saponins **56** and **64** are identical in mass and fragmentation pattern and it was not possible to confirm which isomer corresponds to saponin **K**.

S. brachiatum- 18 peaks between 32 and 61 minutes were detected in our LC-ESI-MS analysis of *S. brachiatum.* Of the *S. radula* standards, saponins **A** and **F** were present in our analysis of the *S. brachiatum* extract. In addition, compound **47** and **93** displayed a parent ion of *m/z* 793, which matches the molecular weight of saponin **O** previously reported from *S. perfoliatum*. However, compound **93** does not possess the same sugar residues based on the observed fragmentation patterns in our ESI-MS studies. The same is true for a compound **73** with the *m/z* 779, which showed an ion fragment at *m/z* 647, corresponding to the loss of a rhamnose [M-146-H]- and thus, did not correlate with the structure of saponin **K,** originally reported from *S. perfoliatum*. Therefore, only compound **47** and **102** could be tentatively assigned as saponins **O** and **N** in our LC-MS analysis of *S. brachiatum.*

The *S. brachiatum* extract had quantitatively the lowest concentration of saponins per gram of dry weight of all species examined in our study. All peaks in the LC-MS analysis of *S. brachiatum* extract were attributed to saponins, whereas, the most abundant

ion peaks in extracts of *S. laciniatum* and *S. terebinthinaceum,* which also had very low concentrations of saponins per gram of dry weight, corresponded to flavonoid glycosides eluting around 25 minutes. Thus, despite the dry weights of extracts, *S brachiatum* probably contains relatively more saponins and in greater quantity than either *S. laciniatum* or *S. terebinthinaceum.*

S. compositum- The *S. compositum* extract analyzed in our LC-MS analysis exhibited 16 peaks between the retention times of 31-59 minutes. *S. compositum* extracts contained saponins with the highest molecular weights (>1000) eluting between 31 and 39 minutes. None of the *S. radula* standards were present in *S. compositum* extracts. However, compound **53** shared the same parent ion as saponin **I**, but differed in retention time and fragmentation pattern. Compound **102**, was detected in the *S. compositum* extract and could be tentatively identified as saponin **N**, reported from *S. perfoliatum*.

S. laciniatum- 9 peaks between the retention times of 32 and 61 minutes were detected in the *S. laciniatum* extract. However, the peaks represent only traces of saponins and the majority of the dry weight of the extract was attributed to a prominent molecular ion of *m/z* 477 at 25.53 minutes, that we tentatively assign as an isorhamnetin monoside, based on mass spectral characteristics, retention times and comparison with literature data (Williams, 2006). Of the trace amounts of saponins found in our LC-MS analysis the *S. laciniatum* extract, none corresponded to the *S. radula* standards. Compound **47** (*m/z* 793) could be tentatively assigned to saponin **O** previously isolated from *S. perfoliatum*,

based on the sequential loss of a hexose moiety [M-162-H]- and glucuronic acid [M-162-176-H]-. Compound **73** (*m/z* 779) shared the same molecular ion as saponin **K**, but differed in retention time and ion fragments and therefore did not correlate with our tentative assignment of this saponin. Compound **102** was detected in the *S. laciniatum* extract and could be tentatively identified as saponin **N**, reported from *S. perfoliatum*.

S. morhii- Our analysis of the *S. morhii* extract showed 17 peaks between the retention times of 27 and 60 minutes, showing the greatest range of retention times among all species tested in this analysis. The *S. morhii* extract did not contain any of the *S. radula* standards. However, Compound **47** and **102**, were tentatively assigned to saponins **O** and **N,** previously isolated from *S. perfoliatum,* based on the identical molecular weights and proposed retention times and fragmentation patterns observed in our LC-MS analysis of the *S. morhii* extract. In addition, we propose compound **39** corresponds to saponin **J**, based on observed fragmentation behavior in this study, previous isolation from *S. perfoliatum* and comparison with literature data (Hattori et al., 1988), (Figure 6.7).

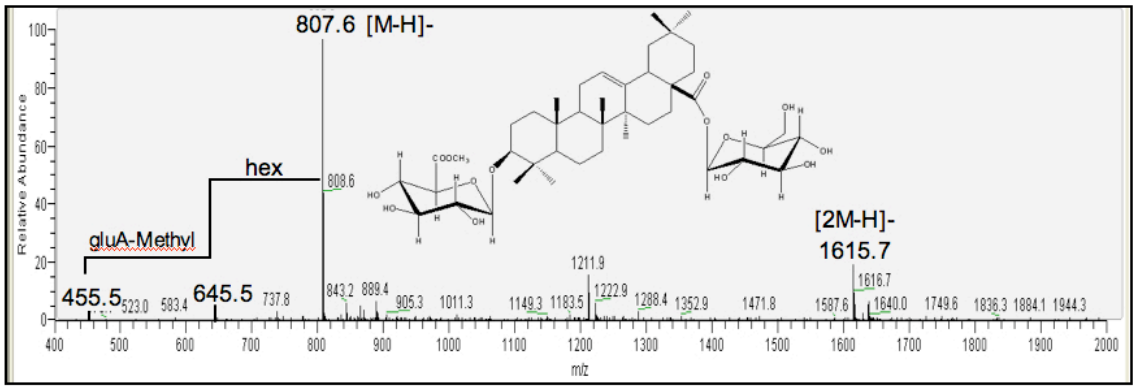


Figure 6.7. Mass spectrum for compound 39, tentatively assigned as saponin J, detected in *Silphium morhii*.

S. terebinthinaceum- Three peaks between the retention times of 46 and 59 minutes were detected in our analysis of the *S. terebinthinaceum* extract. These peaks represent only traces of saponins, quantitatively one of the lowest amounts per gram of dry weight and the fewest number of saponin of all species tested. The majority of the dry weight of the extract was attributed to polar flavonoids that were not removed in the SPE clean-up procedure. These flavonoids were easily distinguished from saponins based on their earlier elution times, mass spectral characteristics and comparison with literature data on flavonoids previously isolated from this species (Williams, 2006). The three saponins that were detected in trace amounts corresponded to compounds **73 (***m/z* 779)**, 87 (***m/z* 647) and **102 (***m/z* 631), respectively. None of the saponin standards isolated from *S. radula* were present in *S. terebinthinaceum*. However, compound **102** could be tentatively assigned to saponin **N**, previously reported from *S. perfoliatum* based on the identical parent ion and additional fragment ions corresponding to the loss of a single glucuronic acid moiety $[M-176-H]^-$. Compound **73** shared the same molecular ion as saponin **K** (*m/z* 779), but differed in the fragment ions.

S. wasiotense- 22 peaks between the retention times of 27 and 60 minutes were detected in the *S. wasiotense* extract, which tied *S. morhii* for the greatest range of retention times of any species. *S. wasiotense* extracts contained one of the standards previously isolated from *S. radula*, saponin **F**. Compound **47** (*m/z* 793), **66** (*m/z* 941) and **102** (*m/z* 631) were tentatively assigned as saponins **O, M** and **N,** reported from *S. perfoliatum*, based on their molecular weights, proposed retention times and mass fragmentation patterns. The

compounds **93** (*m/z* 793) and **4** (*m/z* 941) had the same molecular weights as saponins **O** and **M,** first reported from *S. perfoliatum*, but these compounds were ruled out as positive matches due to their differing mass fragmentations and retention times.

S. integrifolium- 27 peaks between the retention times of 28 and 59 minutes were detected in the *S. integrifolium* extract. It appears that most of these major components occur as isomeric pairs with differing retention times based on their matching mass fingerprints. The majority of the saponins in *S. integrifolium* eluted in a small retention time window between 34 and 48 minutes decreasing the resolution of the chromatogram in comparison with other species examined in our analysis. However, the overall sensitivity of the method combined with extracted ion chromatograms allowed for good overall detection even with peak overlap.

Two of the *S. radula* standards, compound **A** and **D,** were found is extracts of *S. integrifolium*. Compound **56**, **66** and **102** were tentatively assigned to saponin **K** and **M** and **N** based on their matching molecular weights, fragmentation patterns and proposed retention times observed in our LC-MS analysis of the *S. integrifolium* extract. Compound **4** (*m/z* 941) displayed the same parent ion but different fragmentation patterns than saponin **M.** Compounds **48** and **49** (*m/z* 779) also share the same molecular weights and fragmentation as saponin **M**, however these two isomers appear at earlier retention times than the other proposed oleanolic acid saponins present in this study and therefore were ruled out as positive matches to compound **K**.

Lindheimera texana- 18 peaks between retention times of 39 and 60 minutes were detected in the extract of *Lindheimera texana*. Many of the compounds found in *Lindheimera texana* were unique to this species. However, there were some saponins present in *L. texana* that have been previously reported from *S. perfoliatum*; compounds **47** and **102** were tentatively assigned as saponins **O** and **N** based on their identical mass fragmentation patterns. There were also several compounds with matching molecular weights as saponins **K** (**48, 49, 64** and **95***)* and saponin **O** (**93**) present in the *S. wasiotense* extract, but these compounds were ruled out as positive matches due to their differing mass spectral fragmentation and retention times.

6.3.3 Additional Ions of Interest

m/z 809- Compound **13** displayed a molecular ion of *m/z* 809 and showed additional ions at *m/z* 647 and 471, corresponding to the loss of a hexose and a glucuronic acid moiety $[M-162-176-H]^-$ (Figure 6.8). Because there are many aglycones reported in the literature with a molecular weight corresponding to the ion *m/z* 471 and there were no previous reports from *Silphium* of this mass, it was not possible make a tentative assignments of the compound without an unambiguously confirmed standard. However, because Compound **13** was present in seven out of the eleven species of *Silphium*, further investigation of this compound is warranted. (see Table C.1, in the Appendix C, for species names).

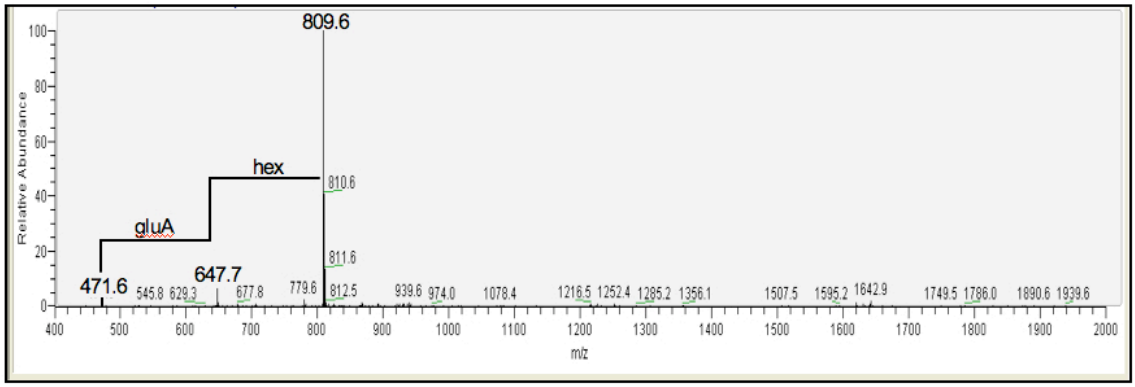


Figure 6.8. Mass spectrum for the unknown compound 13.

In addition to compound **13**, three other saponins were identified with the parent ion *m/z* 809; compounds **1, 25** and **60**. Compound **1** was the earliest eluting saponin in our analyses and was present in two species, *S. wasiotense* and *S. morhii.* Compound **25** was found only in *S. wasiotense* and displayed two additional ions at *m/z* 647 and 485, corresponding to the loss of two hexose moieties with an aglycone of *m/z* 485. Although it was not possible to assign a structure to compound **1** based on the mass spectral data alone, based on the retention times and fragmentation behavior of the polyhydroxylated triterpene saponin standards from *Silphium radula* (Table 6.1), we propose that compound **1** has a structure related to saponin **A**.

Finally, compound **60** was present in four species *(S. albiflorum, S. laciniatum S. morhii and S. wasiotense*) and displayed two prominent fragment ions at *m/z* 663 and 487, corresponding to the loss of a rhamnose and glucuronic acid moiety [M-146-176-H]-. The aglycone showed a further fragment ion at *m/z* 471 resulting from the loss of 16 mu [M-aglyc-OH-H]-.

m/z 647- Compound **87** was detected in nine of the eleven *Silphium* species in our LC-MS analysis and displayed a molecular ion of *m/z* 647 and an additional prominent ion at *m/z* 471, corresponding to the loss of one glucuronic acid moiety. Compound **87** is a good candidate for further structural characterization because this compound is present in the majority of *Silphium* species and, to the authors' knowledge, there are no literature reports for a saponin with matching mass spectral data.

Compound **90** was present in four *Silphium* species examined and displayed an identical mass fingerprint detected for compound **87**. We propose that the two saponins are isomers differing only in the aglycone structure, for example an ursane derivative with a 3-β-glucuronic acid eluting earlier and the oleanane derivative eluting later. This conclusion is supported by other reports in the literature for oleanolic and ursolic acid saponin pairs that display the same order of elution in reverse phase LC-MS analysis as observed here (Martinet et al., 2001), including the isomer pair present in *S. radula* that differ only in their aglycone structures, either ursane (saponin **C**) or oleanane (saponin **E**).

Compound **45** detected in *Silphium morhii* and *S. wasiotense*, also displayed the same mass fingerprint as compound **87** and **90**, but eluted much earlier than the latter two compounds, suggesting **45** may possess a more polar aglycone.

m/z 925- The literature reports saponins **J, K, L** and **O** (also known as silphiosides B, C, E and G**)** as the major extractable saponins for *S. perfoliatum* (Davidyants et al., 1984a; Davidyants et al., 1984b). However, our results found that a saponin with the molecular

ion of 925.6 *m/z* (**24**) in negative mode was the most abundant saponin in the extract of *S. perfoliatum.* Figure 6.9 presents the mass spectra for compound **24**. This compound was present in several other species including: *S. compositum, S. albiflorum, S laciniatum, S. wasiotense, S brachiatum* and *S. asteriscus*. The loss of a single hexose unit was evident by the strong molecular ion of 763 *m/z* [M-162-H]-. A saponin of the same molecular weight whose structure was determined to be 3-O-β-glucuronopyranosyl-oleanolic acid 28-O-β-glucopyranosyl 1-2- arabinopyranosyl was previously isolated from *Panax japonicus* with the trivial name, chikusetsusaponin IV (Lin, 1976). In addition, the methyl ester of this compound, known as chikusetsusaponin IVa, was also isolated from *Panax japonicus* and *Silphium perfoliatum* (Lin, 1976; Davidyants et al., 1984a,b). Thus, we tentatively assigned compound **24** as chikusetusaponin IV based on the identical MS spectra and on the presence of this compound with several structurally related glucuronide-type saponins (chikusetusaponin IVa, V) in *Silphium perfoliatum.*

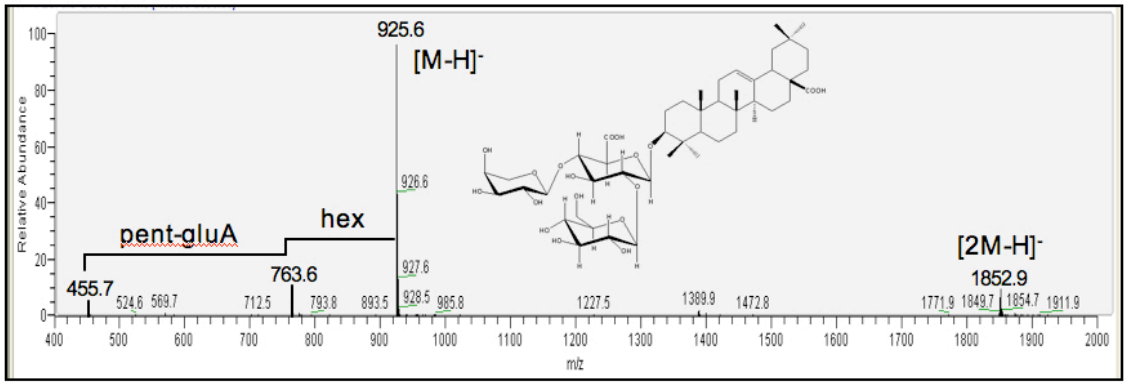


Figure 6.9. Mass spectrum for compound 24. Based on the matching spectra and the presence of this compound and several structurally related glucuronide-type saponins (chikusetusaponin IVa, V) in *Silphium* species, we propose this compound is chikusetsusaponin IV.

***m/z* 955-** Compound **27, 36** and **76** have a molecular ion of *m/z* 955. The product ion spectra for compound **27** showed an additional ion at *m/z* 793 attributed to the loss of one hexose residue. Compound **36** showed additional ions at *m/z* 793, 617 and 455, corresponding to the loss of two glucose and one glucuronic acid moiety. This fragmentation pattern matches that of glycoside B isolated from *Calendula officinalis*, also a member of the Asteraceae (Vidal-Ollivier et al., 1989). Compound **76** showed ion fragments at *m/z* 793, 631 representing the lost of two hexose moieties.

***m/z* 779-** Two compounds (**65** and **74)** with a molecular weight of *m/z* 779 were observed with differing retention times (Rt= 43.21-44.65 and 44.59-45-99)**.** The fragmentation patterns observed for compound **65** and **74** exhibited the loss of two hexose units, suggesting that the pair differ only in the sugar linkage (monodesmoside vs. bidesmoside) or conformation. Huhman et al. (2002) showed that 3-Glc-28-Glc-medicagenic acid conjugates eluted significantly earlier than the 3-Glc-Glc-medicagenic acid conjugates in reverse phase separations with an acid modifier. This difference is due to ability of the carbohydrates on the bidesmoside to minimize van der Waal interactions with the stationary phase, therefore lowering the retention time (Huhman, 2002). A much greater difference in retention time would be observed in isomeric pairs where one is a bidesmoside and the other a monodesmoside. Thus, we propose the earlier eluting isomer pair (**56** and **57**) present in *S. perfoliatum, S integrifolium, S asteriscus* and *S. radula* corresponds to the bidesmoside isomers (3-hex-28-hex-455), while the latter eluting pair (**66** and **75**) present in *S. radula* and *S asteriscus* corresponds to the monodesmoside isomers (3-hex-hex-455). It is not possible to differentiate between the type of hexose

(glucose or galactose) present using ESI-MS alone, therefore only tentative assignments could be made. Compound K (silphioside B) has a structure of 3-O-Glc-28-glu-oleanolic acid (779 *m/z*), thus we propose that compound **56** corresponds to saponin **K**.

6.3.4 Saponins from Kowalski's (2007) LC-MS Studies of *Silphium* Species

Kowalski (2007) recently published a qualitative and quantitative evaluation of triterpene aglycones from saponin fractions of three species of *Silphium, S. perfoliatum, S. trifoliatum* and *S. integrifolium*, as compared *with Panax quinquefolium* root and *Calendula officinalis* flowers using HPLC-PDA/ESI/MS analyses. The results from this study revealed the presence of 36 oleanolic and ursolic acid saponins in leaves, stems and roots of the *Silphium* species examined. In addition, five saponins reported from *Calendula officinalis* were also present in Kowalski's analysis of extracts of *Silphium*, including two previously reported glucuronide saponins, glycoside D2 (also known as calenduloside F) (*m/z* 793) and glycoside F (also known as calenduloside E) (*m/z* 631) (Vidal-Ollivier et al., 1989). In addition to *Calendula officinalis,* these two saponins were also previously isolated from *Silphium perfoliatum* (Figure 6.2; Saponin **N** and **O**) and besides glucuronide D2 and F, no authentic standards were available for Kowalski's (2007) analysis. Therefore, no tentative assignments of the observed ions were made in the HPLC-PDA/MS screening of *Silphium* species.

Table 6.3 compares the saponins detected in the current LC-MS analysis with Kowalski's (2007) analysis of saponins from *Silphium* species. Three saponins with the molecular ion *m/z* 793 were detected in Kowalski's analysis of *S. perfoliatum, S.*

trifoliatum (*S. asteriscus*) and *S. integrifolium* (Kowalski, 2007). The earliest eluting saponin was detected in *S. perfoliatum* leaves and *S. integrifolium* and *S. trifoliatum* rhizomes and roots. According to Kowalski (2007), this saponin showed an identical retention time to glucuronide D2 previously isolated from *Calendula officinalis* (Vidal-Ollivier et al., 1989). Compound **47** in our analysis was detected in several species (see Table C.1 in Appendix C for details) and the fragmentation patterns observed were identical to those reported for glucuronide D2 from *Calendula officinalis* (Vidal-Ollivier et al., 1989). Based on these combined data, we proposed that compound **47** is glucuronide D2.

m/z	**Column A Compound # Present analysis**	**Species**	**Column B Compound # Kowalski (2007)**	**Species**
955	27, 36, 76	PER, ALB, LAC, WAS, BRA, LIN	13	PER, TRI, INT
971	9, 14	WAS, INT	16, 18	PER, TRI, INT
793	47 (Saponin O)	ALB, WAS, BRA, LIN, TER, LAC, MOR, AST	20, 29	PER, TRI, INT
881	63	INT	28, 30, 33, 34	PER, TRI, INT
631	102 (Saponin N)	All species	35	PER, TRI, INT

Species names are abbreviated as follows: ALB= *S. abliflorum*, AST= *S. asteriscus*, BRA= *S. brachiatum*, INT= *S. integrifolium*, LAC= *S. laciniatum*, LIN= *Lindheimera texana*, MOR= *S. morhii*, TER= *S. terebinthinaceum*, WAS= *S. wasiotense*. In addition, species TRI (*S. trifoliatum*) in Kowalski's (2007) analysis is referred to as *S. asteriscus* in our analysis.

Table 6.3. Table comparing saponins detected in extracts of *Silphium* species (A) the current LC-MS analysis with (B) Kowalski's (2007) analysis. In Kowalski's (2007) analysis, roots, rhizomes, inflorescences and leaves were tested. Information on saponin content of the different plant organs are not provided in this table for simplification.

6.4 Conclusions

The HPLC-ESI/MS method developed for the purposes of this study allowed for a rapid and reliable qualitative determination of over 100 saponins in extracts of all *Silphium* species and *Lindheimera texana*. The results from our analysis indicate a much broader range of saponins than were originally reported from any *Silphium* species studied to date (Calabria et al., 2008; Kowalski, 2007). In addition to establishing the presence of previously identified saponins from *S. radula* in other species of *Silphium* by comparison with authentic standards, we putatively identified a large number of triterpene saponins from all other species examined in this study for the first time that would not have been possible using the relatively time consuming isolation and structure determination procedures typically employed for saponins.

Although *Silphium* has been recognized for centuries as a folk medicine, to date there are only a few reports outlining the bioactivities of saponins isolated from *S. radula* and *S. perfoliatum* (Calabria et al, 2008; Syrov et al., 1992; Davidyants et al., 1997). The data from our HPLC-MS analysis have shown that these medicinally important compounds are present in other *Silphium* species (see Table C.1, Appendix C for details). In addition, the tentative identification of other bioactive saponins, not previously reported from *Silphium*, adds to the potential of this genus as a rich source of medicinally important compounds. For example, we tentatively identified the most abundant compound (*m/z* 925) in *S. perfoliatum* as chikusetsusaponin IV, first reported from *Panax japonicus*, which has recently shown to exhibit significant anti-obesity action in rats (Han et al, 2005). Obesity is the second leading cause of preventable death in the United States

(Mokdad et al., 2004). Furthermore, saponins from *Calendula officinalis*, which were detected in our current analysis of *Silphium* species (saponin **N** and **O)**, have well-established pharmacological activities, such as hypoglycemic, gastroprotective, antiviral, anti-mutagenic and anti-inflammatory (Ukiya et al., 2001; Ukiya et al., 2002; Yoshikawa et al., 2001).

Of all the saponins detected in our LC-MS analysis of *Silphium* and *Lindheimera texana*, only one (compound **102**) was present in all species. This compound was tentatively identified as saponin **N,** previously reported from *S. perfoliatum* and *Calendula officinalis* and has been shown to exhibit both fungicidal and anti-bacterial properties (Szakiel et al., 2005). The constitutive expression of saponin **N** in all extracts tested in our LC-MS analysis suggests that this compound may be widespread in the Asteraceae family.

It is likely that the great variation in saponin profiles observed between species of *Silphium* represent a combination of environmental, developmental and genetic factors. Population level chemical screenings would greatly aid in understanding which of these factors contribute most to saponin variation in *Silphium*. The genus-wide metabolic profiling of saponins presented here provides a foundation for more comprehensive surveys of saponin distribution in the Asteraceae, and more specifically, in closely related genera in the Engelmanniinae subtribe and Heliantheae tribe. These types of chemical inventories are necessary to understand the distribution and functional significance of saponins in plants and ultimately for future efforts to promote the health benefits of these medicinally important compounds.

Chapter 7: Conclusions

This dissertation describes phytochemical and chemosystematic studies of one of the largest and most diverse plant families in the world, the Sunflowers. More specifically, most of my doctoral research has focused on the triterpene saponin chemistry of *Silphium* L. (Asteraceae), a small genus of perennials sunflowers native to North America and naturalized to parts of Europe where it is cultivated as a garden ornamental and herbal medicine.

The primary objective of my dissertation research was to develop and execute a standardized procedure for the extraction and identification of saponins from *Silphium* species in order to determine the precise chemical structures of these compounds. In the course of these investigations, nine new triterpene saponins were identified from *Silphium radula* and one from *S. integrifolium*. Several other known saponins were also isolated and identified from *Silphium integrifolium* and *S. morhii*. The structures of these compounds were established by chemical methods and spectral analyses including ^{1}H-NMR, ^{13}C NMR, HMBC, NOESY DEPT and TOSCY experiments, as well as high resolution and low-resolution ESI-MS analyses. The new compounds were identified as 3β,6β,16β-trihydroxyolean-12-en-23-al-3-*O*-β-glucopyranosyl-16-*O*-β-glucopyranoside (**1**), urs-12-ene-3β,6β,16β-triol-3-*O*-β-galactopyranosyl-(1 → 2)-β-glucopyranoside (**2**), 3β,6β,16β-trihydroxyolean-12-en-23-oic acid-3-*O*-β-glucopyranosyl-16-*O*-β-glucopyranoside (**3**), urs-12-ene-3β,6β,16β,21β-tetraol-3-*O*-β-glucopyranoside (**4**), olean-12-ene-3β,6β,16β,21β-tetraol-3-*O*-β-glucopyranoside (**5**), olean-12-ene-3β,6β,16β,21β,23-pentaol-3-*O*-β-glucopyranosyl-16-*O*-β-glucopyranoside (**6**), olean-12-

ene-3β,6β,16β-triol-3-*O*-β-glucopyranosyl-16-*O*-α-arabinopyranosyl-(1 → 2)-β-glucopyranoside (**7**), olean-12-ene-3β,6β,16β,23-tetraol-3-*O*-β-glucopyranosyl-16-*O*-α-arabinopyranosyl-(1 → 2)-β-glucopyranoside (**8**), 3β,6β,16β,21β-tetrahydroxyolean-12-en-23-al-3-*O*-β-glucopyranoside (**9**) and ursolic acid 3-*O*-β-glucuronopyranosy-28-*O*-β-glucopyranoside (**10**). Moreover, the cytotoxic activities of the isolated compounds were tested against human breast cancer cell line MDA-MB-231. Results showed that compound **2** decreased cell proliferation in a statistically significant manner at concentrations of 25 µg/ml.

A second major objective of my research was to establish a comprehensive HPLC/ESI-MS method for the metabolic profiling of saponins in *Silphium*. This genus-wide analysis of *Silphium* leaf extracts revealed the presence of over 90 additional saponins in the *Silphium* genus and in the closely related *Lindheimera texana*. This methodology allowed for the putative identification of a large number of triterpene saponins from all species examined that would not have been possible using traditional methods. In addition, the data produced from these analyses allowed us to examine the chemosystematic value of saponins in these taxa at the genera and species level. Our studies provided additional reports of saponins for eleven species in the Asteraceae including new reports for saponins from nine species of *Silphium* and for *Lindheimera texana*. These results indicate that the Engelmanniinae subtribe, and perhaps many other Asteraceae are a rich source of new saponin structures.

As a result of my investigations of triterpene saponins from *Silphium,* I became interested in the chemosystematic value of these compounds for understanding

evolutionary and systematic aspects of the Asteraceae. One way to evaluate the chemosystematic potential of secondary metabolites for a particular group of plants is to use these compounds as chemical characters to construct phylogenetic trees. Thus, with my colleagues, Prof. Emerenciano and his graduate student Marcus Scotti, the first phylogenetic analysis of the entire Asteraceae based only on chemical data was conducted. The data matrix used in this study was created from a large chemical database comprising ~400 skeletal types of terpenes, coumarins, flavonoids, benzofurans and polyacetylenes isolated from the family (Gastmans et al., 1990a and 1990b). Hypotheses about the relationships among tribes were discussed and compared with relationships inferred from the morphological and macromolecular classifications. The results from this analysis indicated that the phylogenetic signal of the chemical data is compromised by the incomplete and often inconsistent chemical reports and until this problem with the literature data is overcome, molecular data will continue to provide the most reliable phylogenetic trees. However, DNA data alone cannot explain how or why a particular plant evolved and a single view using one type of approach, whether it be molecular, morphological or chemical, cannot provide answers to all questions relating to the evolution of the Asteraceae.

Perhaps a more complementary way to utilize phytochemical data for systematic purposes is to map the current distribution, abundance, and diversity of secondary metabolites taken from an extensive chemical database for the Asteraceae onto the most current DNA-based phylogenetic trees for the family. Our results indicate that there are thousands of secondary metabolites described from nearly every tribe of the Asteraceae.

Many interesting patterns emerged from viewing the chemical data within a well-supported DNA-based phylogenetic framework. Still, the majority of the literature data used in our meta-analysis were not generated for taxonomic purposes and the biological activities of the reported chemicals were greatly lacking in comparison to the structural data. Thus, in order to understand the functions of secondary metabolites in plants, chemists need to focus not only on the isolation and characterization of new compounds but also on the analysis of the biological functions of known compounds. Moreover, by generating and interpreting chemical data for taxonomic purposes, researchers who understand the distribution, ecology and molecular phylogenies of a particularly group of plants can use these data to ascertain the evolutionary significance of these compounds for plant survival and adaptation. As we continue to do this, the taxonomic value of these compounds will increase.

Finally, I will mention the importance of training new generations of phytochemists in the skills required for the isolation and characterization of secondary metabolites in plants. The executive summary on phytochemistry in The Report for the Next Millennium (Botanical Society of America, 1995) summarized this need very clearly:

> "….At the moment, there is no set of funding programs at federal or international agencies or at private foundations that caters to education or basic research in phytochemistry. Most phytochemists are unable to do the work they are trained to do because of the lack of research funding, which is correlated with fewer students choosing to be trained in phytochemistry. This means that the would-be pillars of phytochemistry are weakening. Unless priorities in phytochemical research and education are recognized and supported, knowledge about plant chemistry that will be needed in other disciplines in biology will become dated or remain incomplete. It is therefore crucial to fund phytochemical research and to support programs for training professional phytochemists…."

Now, over 10 years from the time that this was written, there are still no major funding programs that support basic research and education in phytochemistry. I was extremely fortunate to be a part of one of the most prestigious and long-standing phytochemical laboratories in the country, which after more than 40 years, will cease to exist upon my graduation. I will officially be the last classically trained phytochemistry student in the Plant Biology program at UT-Austin.

Still, I feel confident that students throughout the plant sciences will continue to be intrigued by the mysteries of plant secondary metabolism. With increasingly sensitive methods for the detection of natural products in plant extracts, it will become easier for researchers to contribute to our understanding of the distribution of phytochemicals from diverse plant sources. Moreover, because of advances in molecular technologies for the rapid analysis of DNA, more accurate phylogenies have become available for examining the chemistry of plants in the context of their life histories. But even with these advancements, the underlying role of secondary metabolites in the evolution and diversification of plants can only be understood through a multi-disciplinary approach. Therefore, the most fundamental understanding that I take away from my studies of the

triterpene saponin chemistry of *Silphium,* a small, inconspicuous genus in the Sunflower family, is the overarching importance of phytochemical research in all of the plant sciences.

Appendix A

.

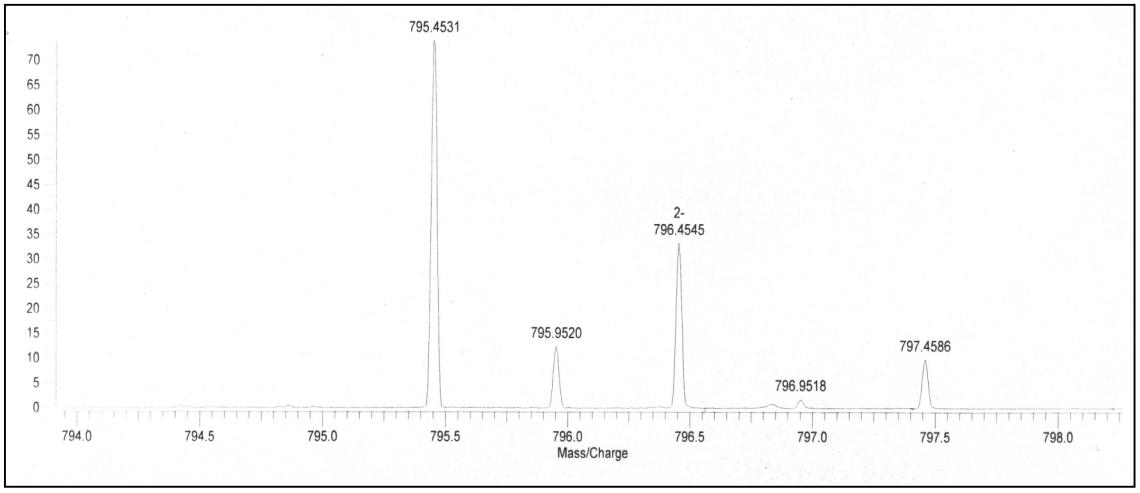


Figure A.1. HR-ESI-MS negative mode spectra for *S. radula* compound **1.**

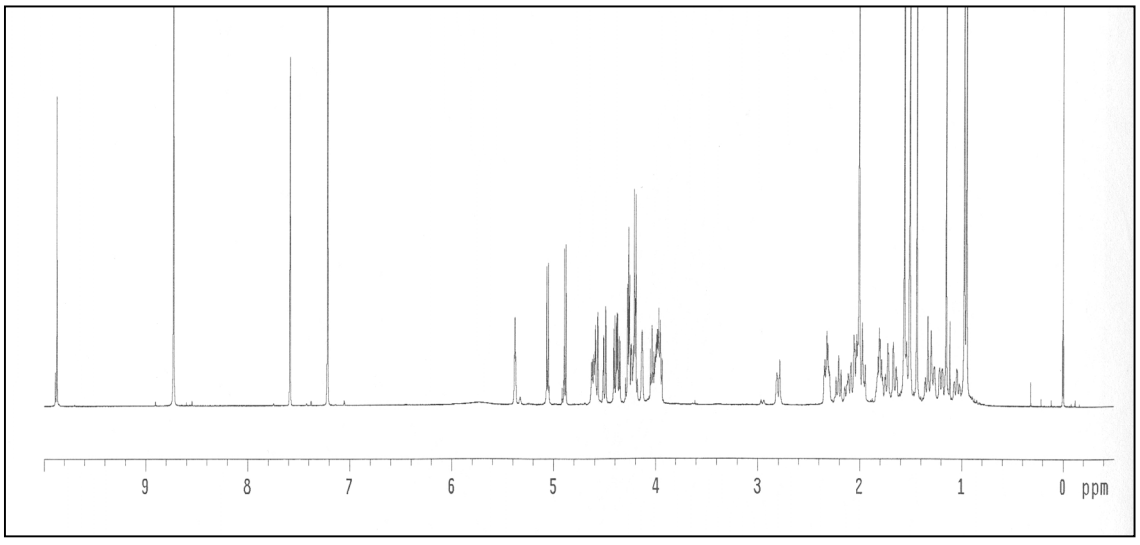


Figure A.2. ^{1}H-NMR spectra for *S. radula* compound **1**.

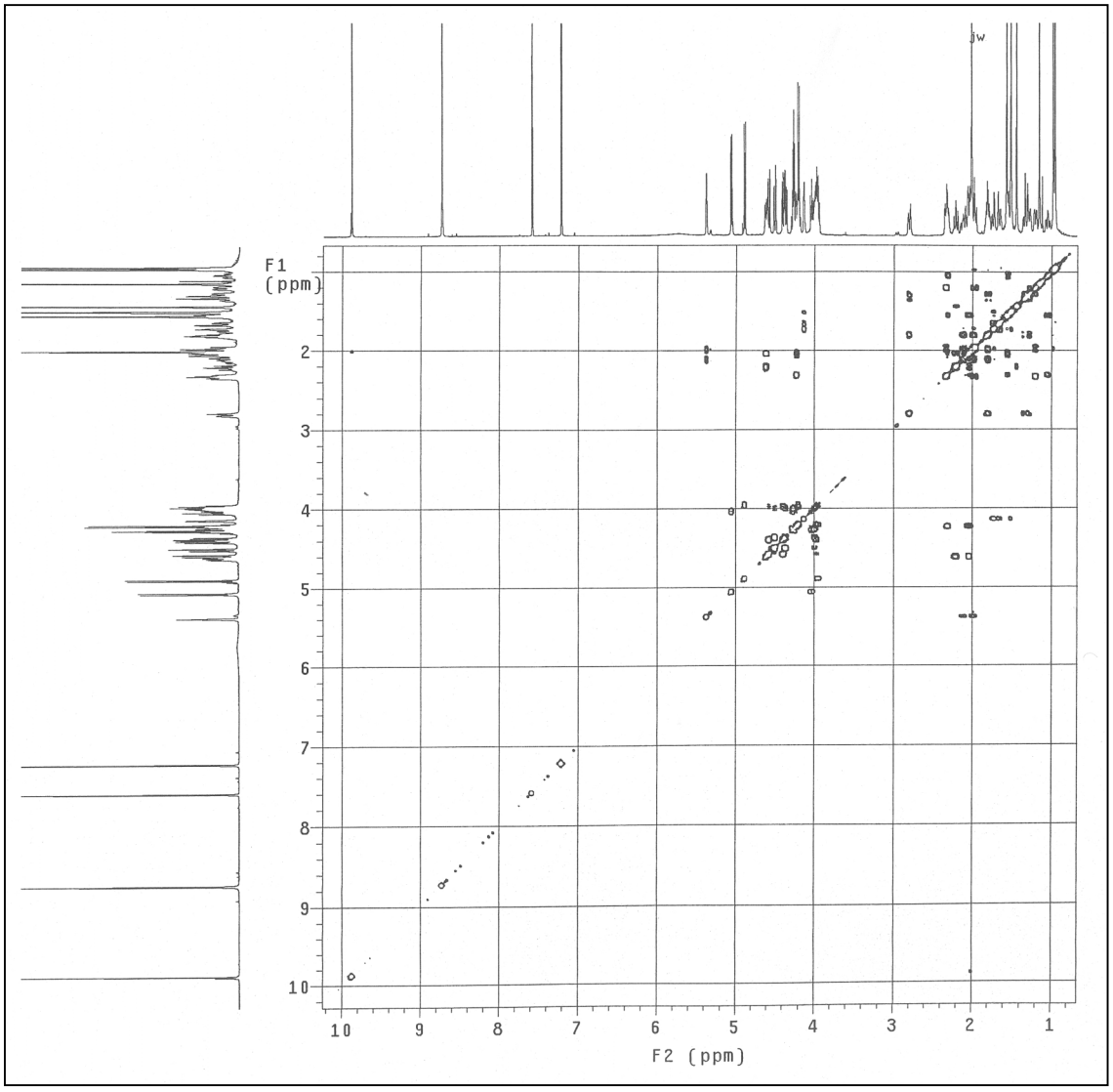


Figure A.3. COSY spectra for *S. radula* compound **1.**

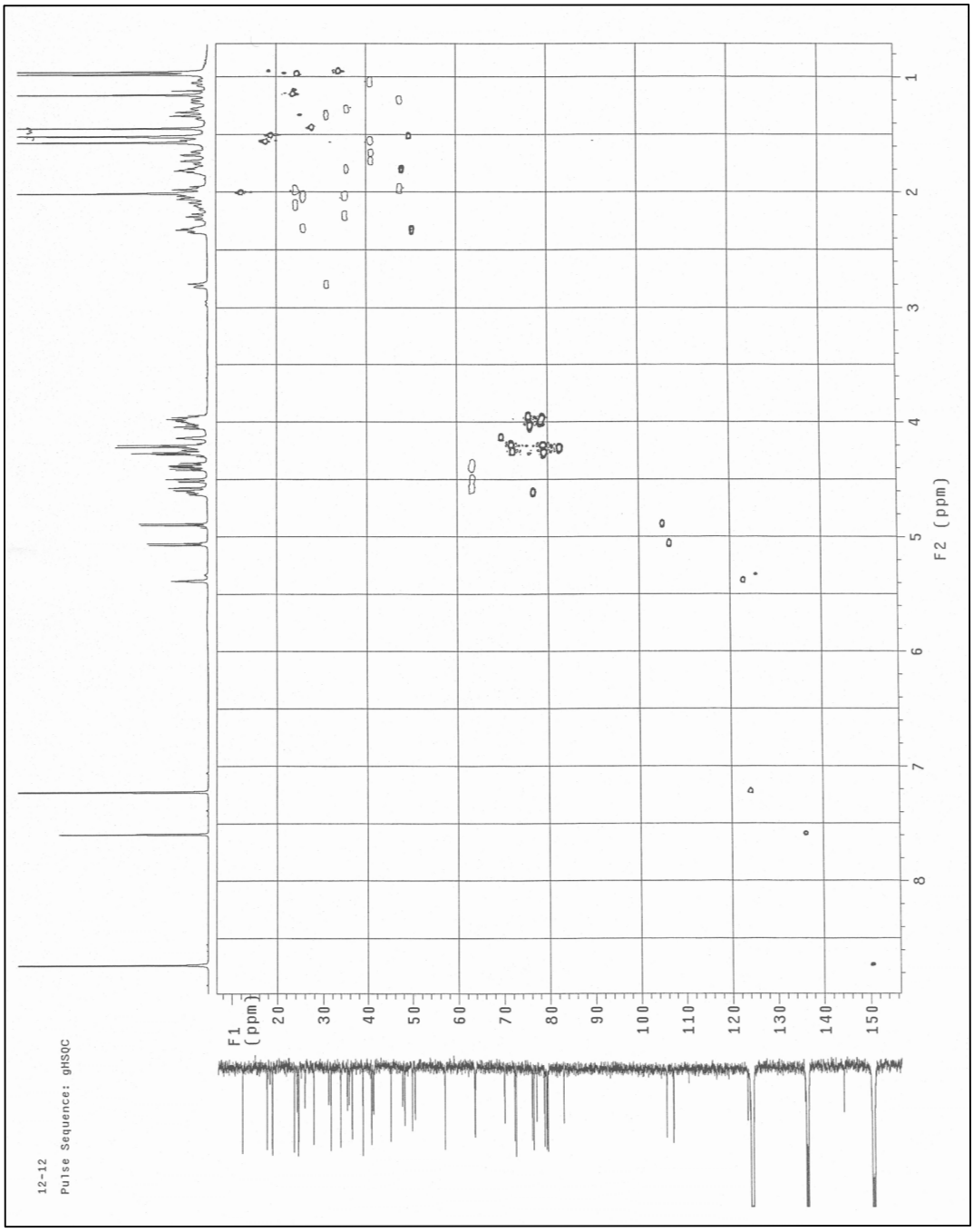


Figure A.4. HSQC spectra for *S. radula* compound **1.**

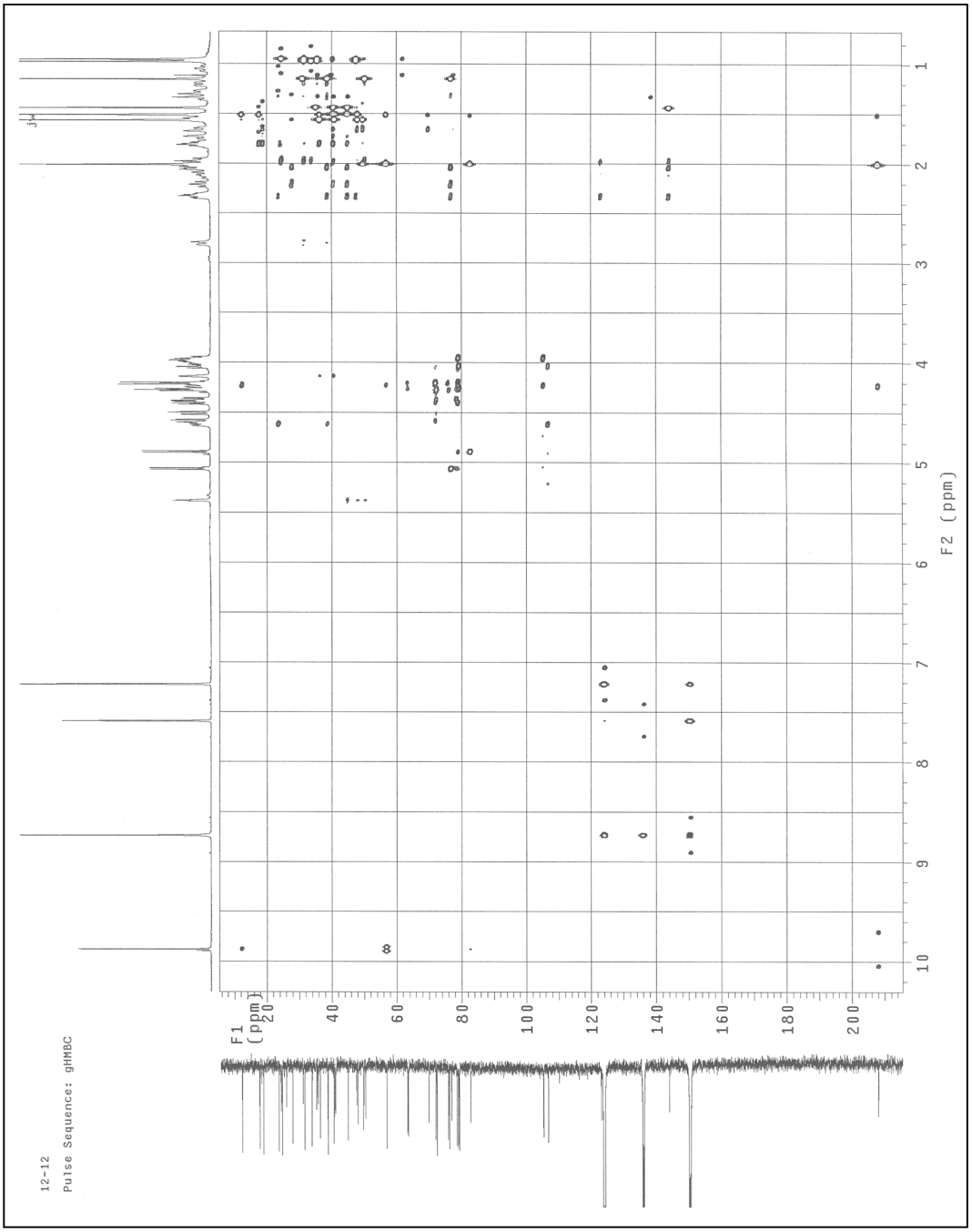


Figure A.5. HMBC spectra for *S. radula* compound **1**.

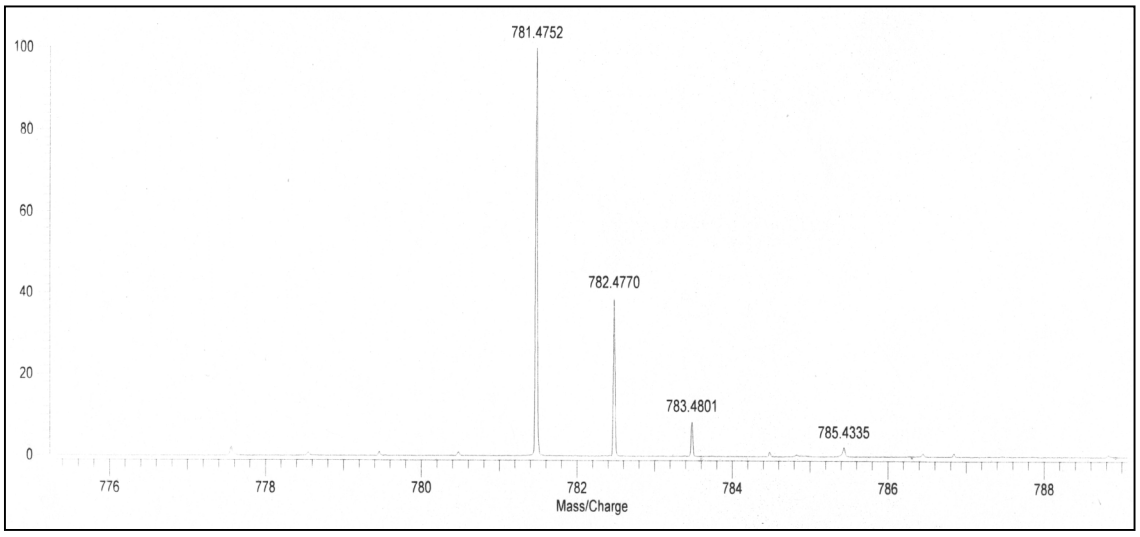


Figure A.6. HR-ESI-MS negative mode spectra for *S. radula* compound **2**.

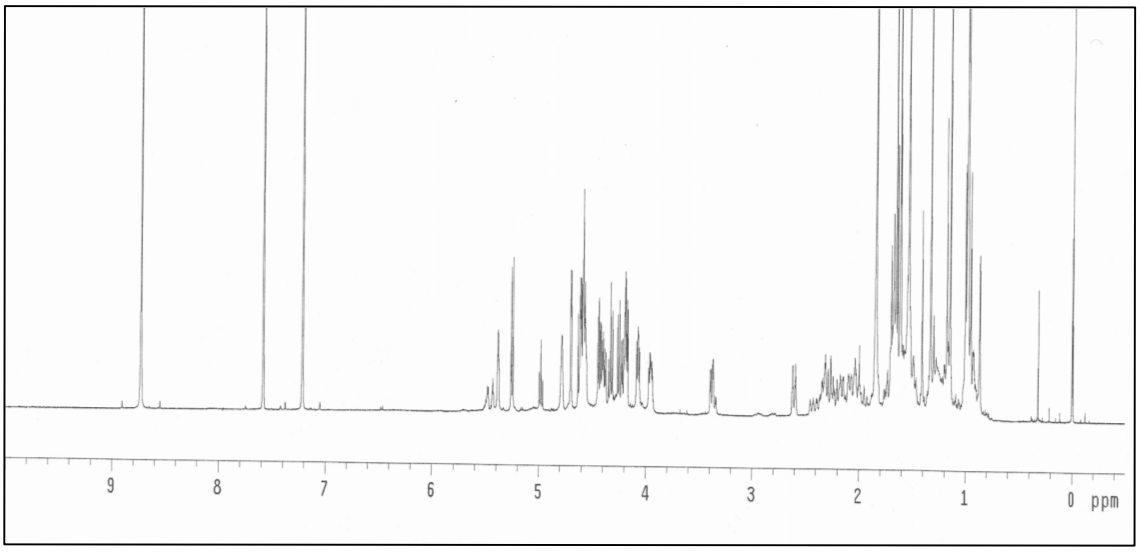


Figure A.7. ^{1}H-NMR spectra for *S. radula* compound **2.**

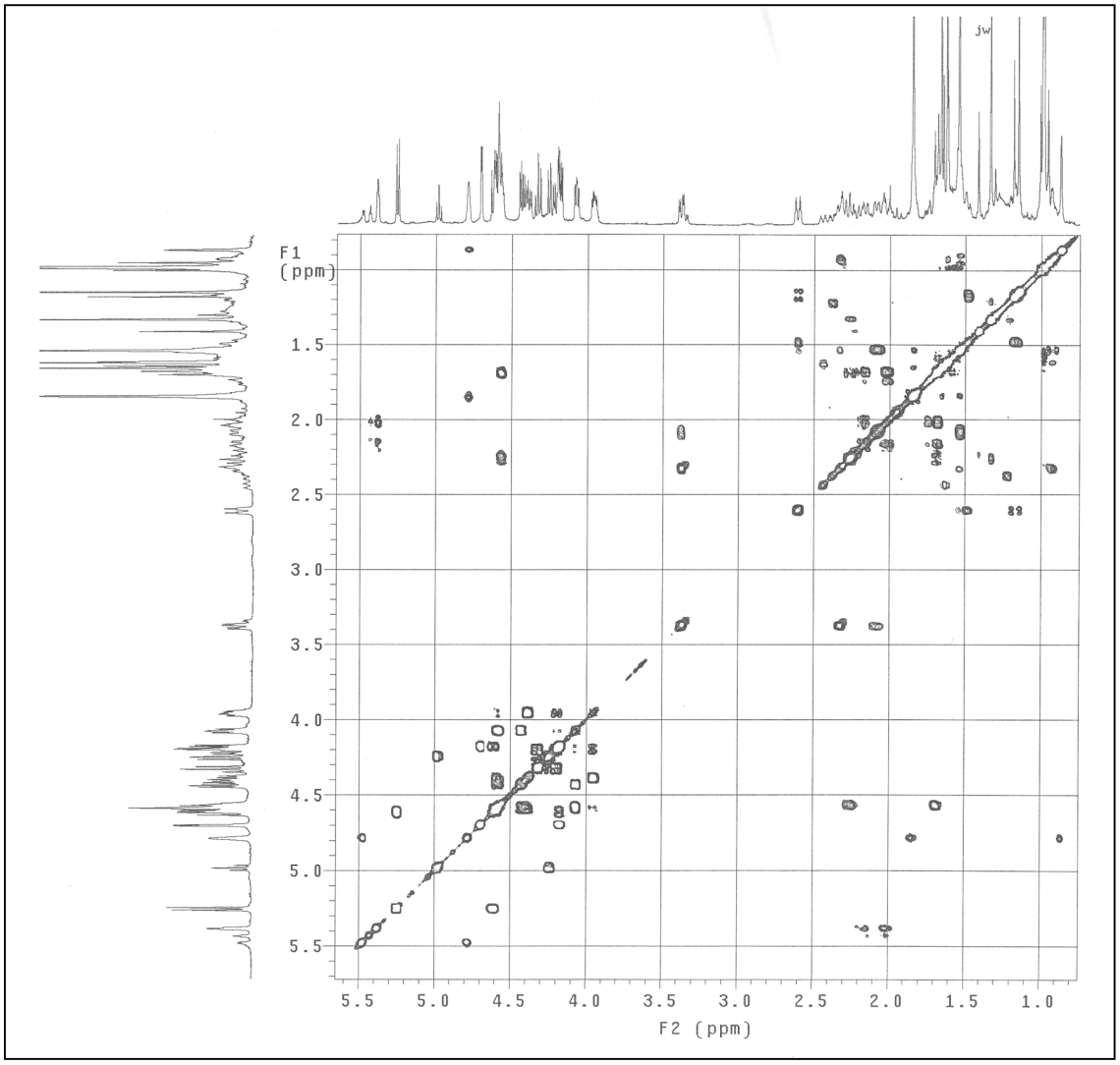


Figure A.8. COSY spectra for *S. radula* compound **2**.

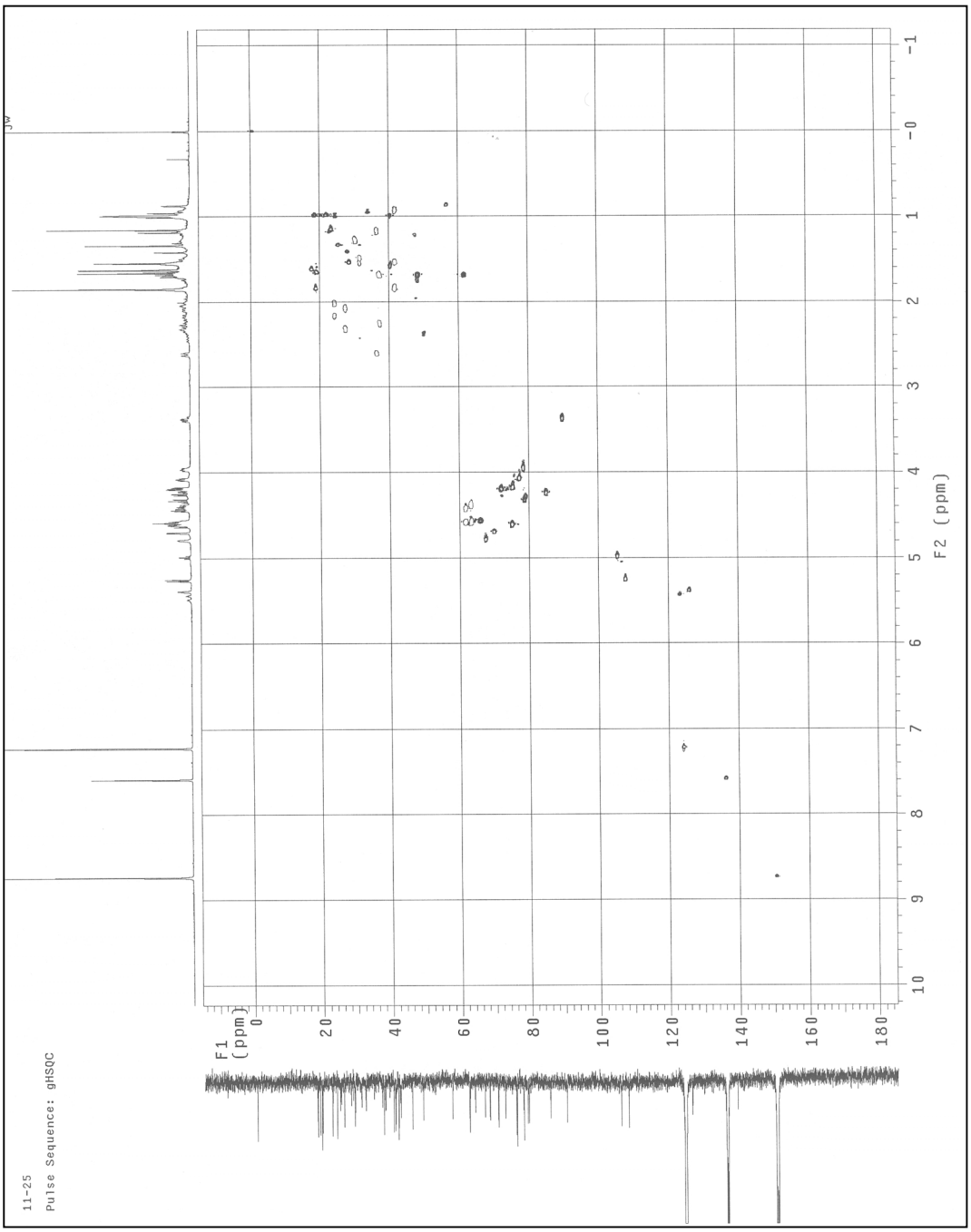


Figure A.9. HSQC spectra for *S. radula* compound **2.**

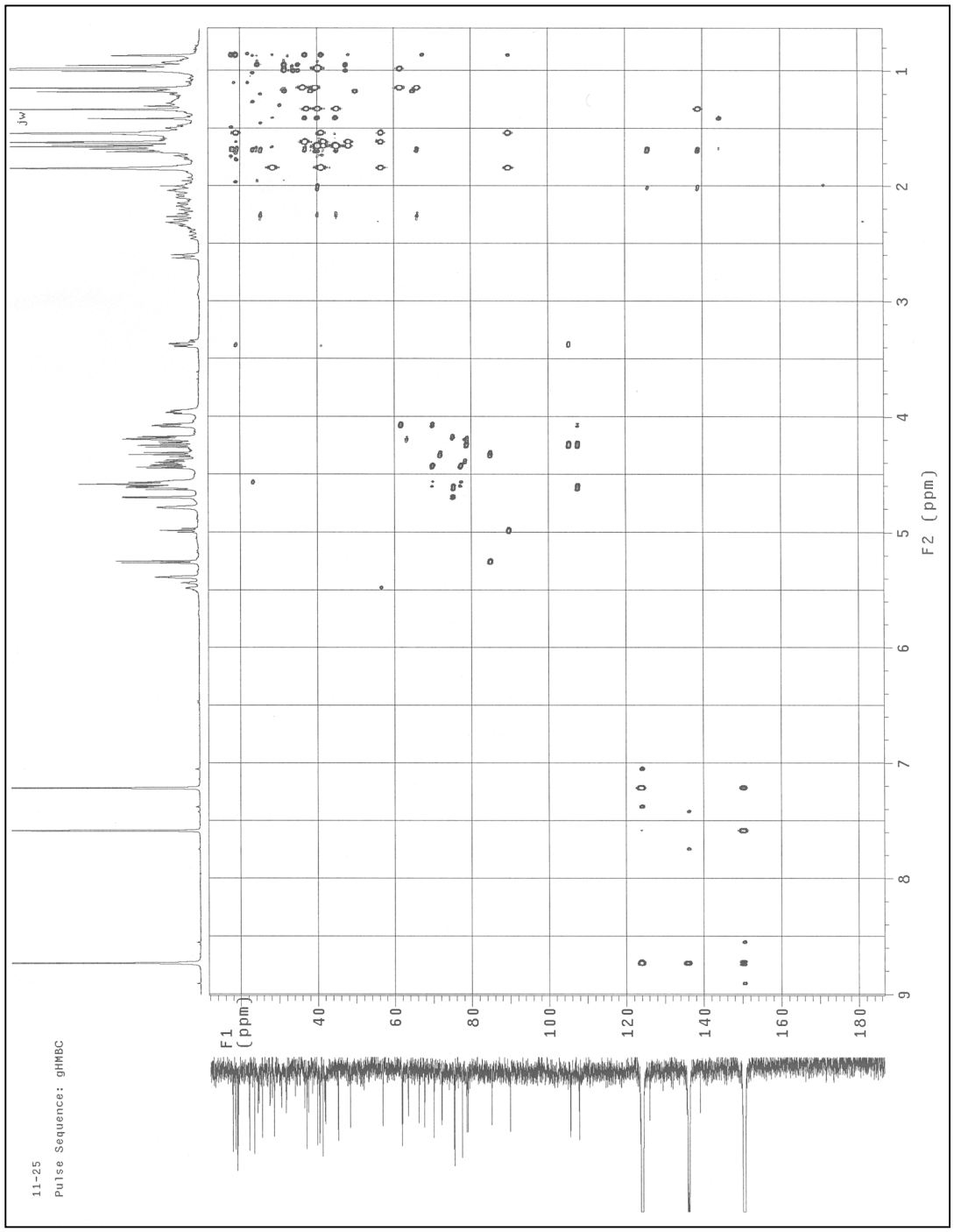


Figure A.10. HMBC spectra for *S. radula* compound **2**.

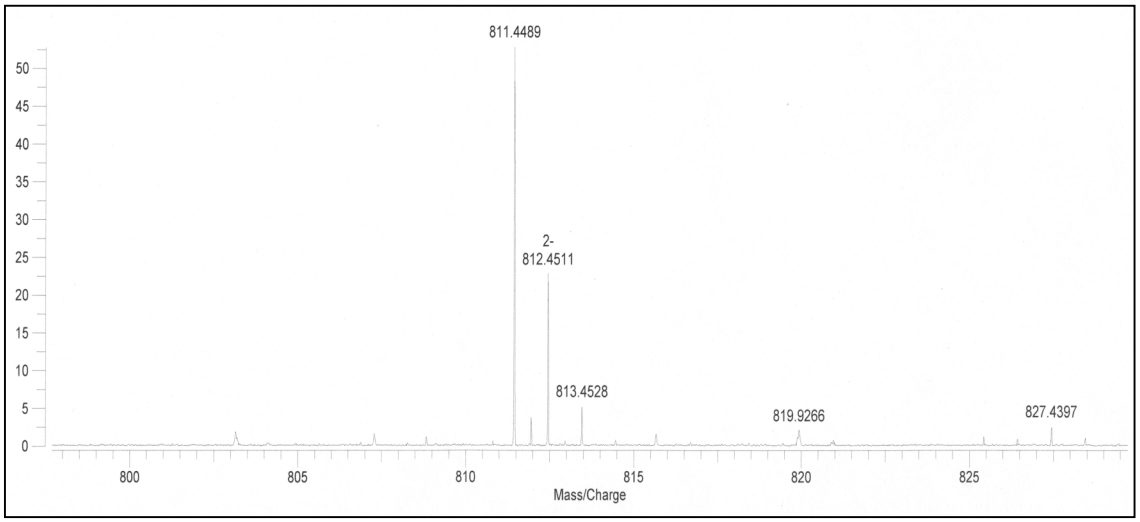


Figure A.11. HR-ESI-MS negative mode spectra for *S. radula* compound **3.**

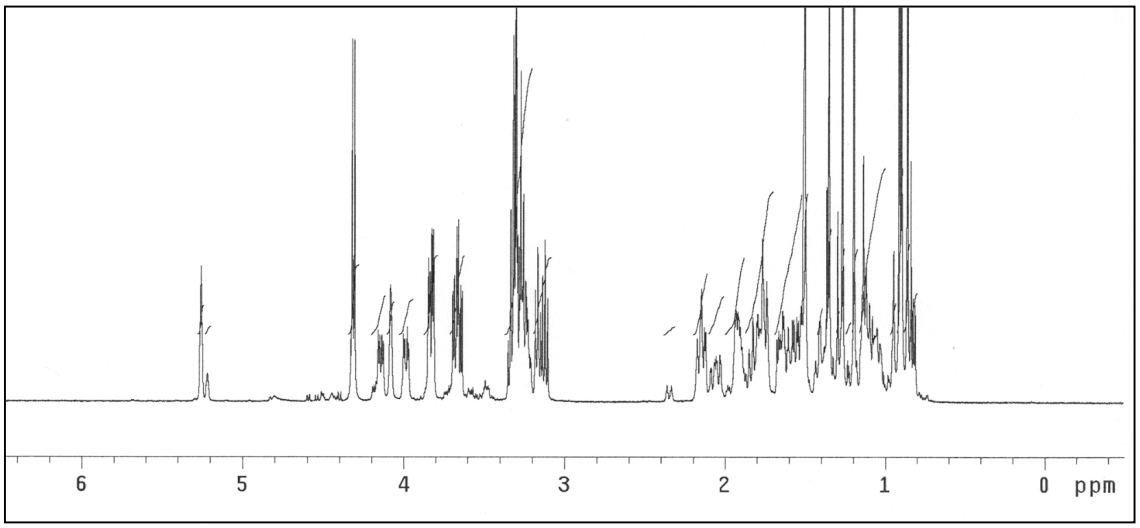


Figure. A.12. ^{1}H-NMR spectra for *S. radula* compound **3**.

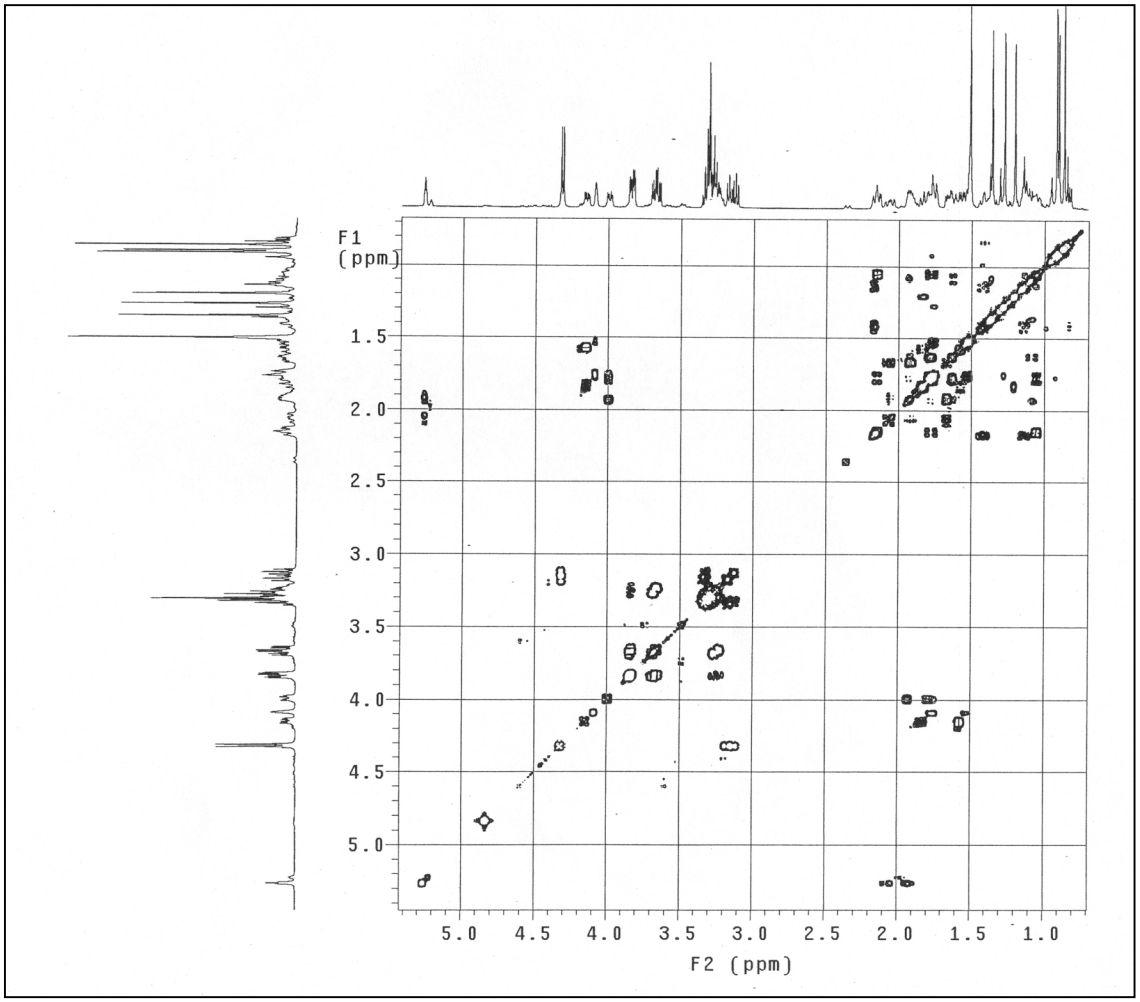


Figure A.13. COSY spectra for *S. radula* compound **3.**

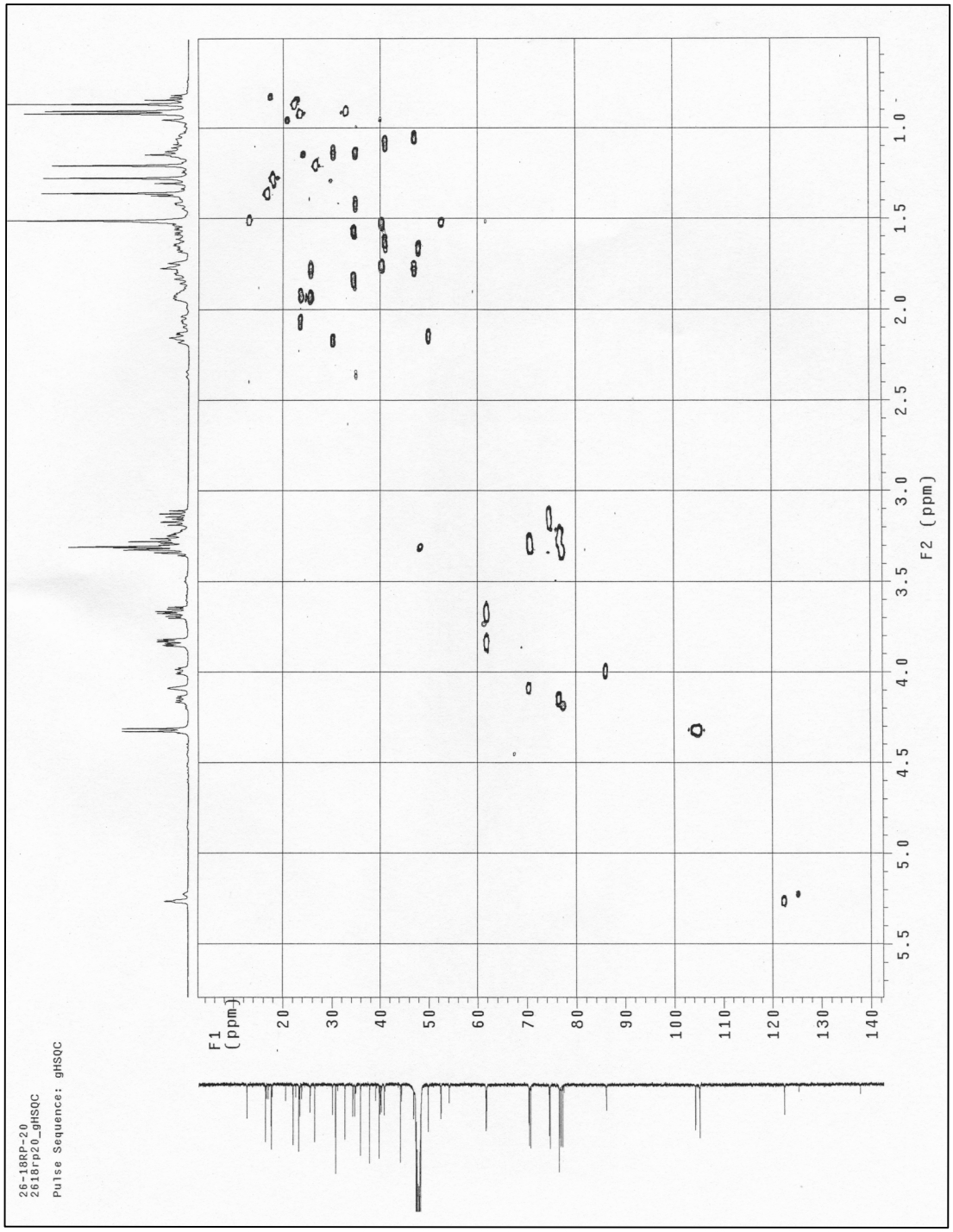


Figure A.14. HSQC spectra for *S. radula* compound **3**.

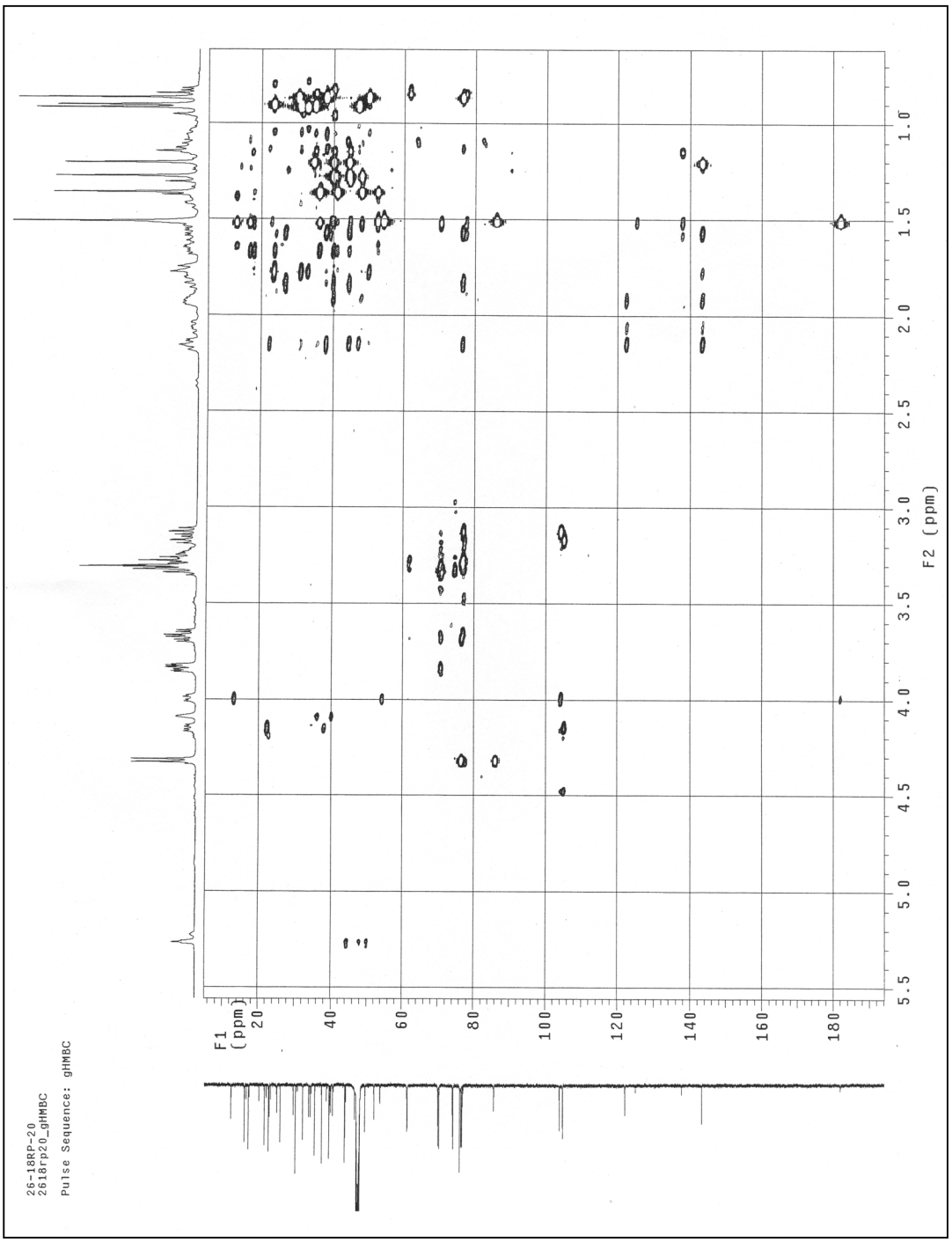


Figure A.15. HMBC spectra for *S. radula* compound **3.**

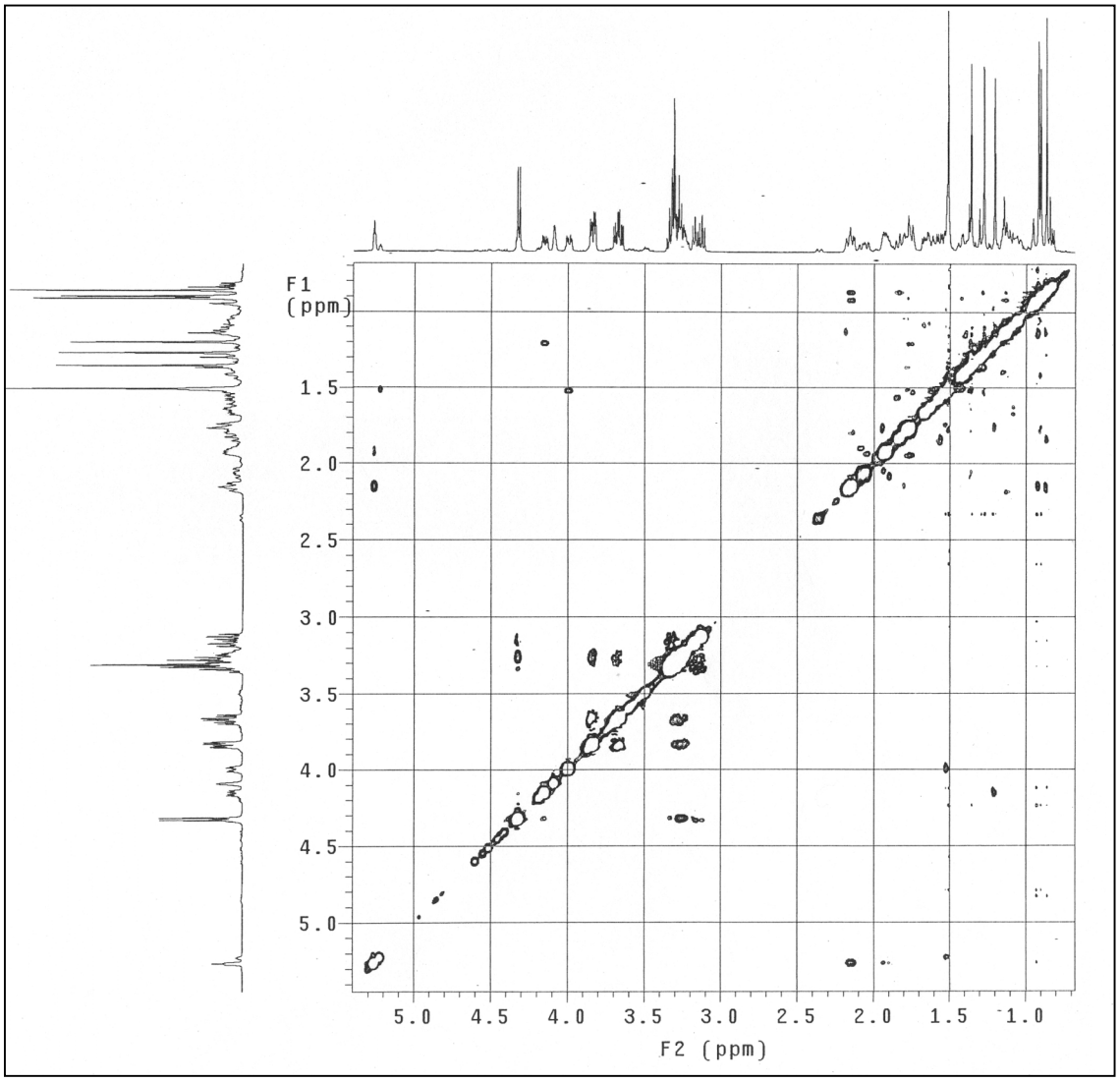


Figure A.16. NOESY spectra for *S. radula* compound **3**.

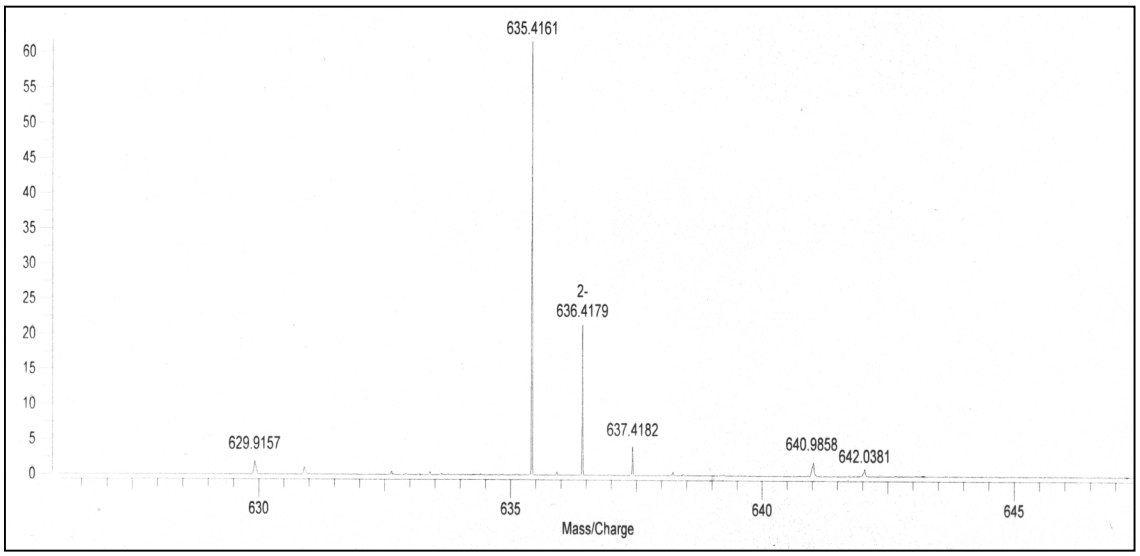


Figure A.17. HR-ESI-MS negative mode spectra for *S. radula* compound **4**.

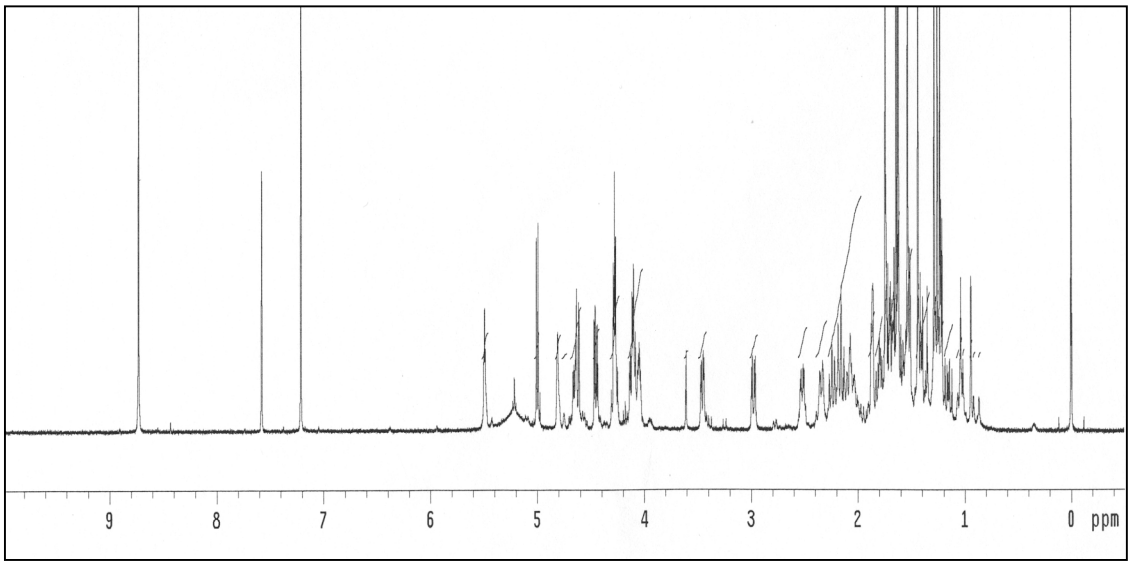


Figure A.18. ^{1}H-NMR spectra for *S. radula* compound **4**.

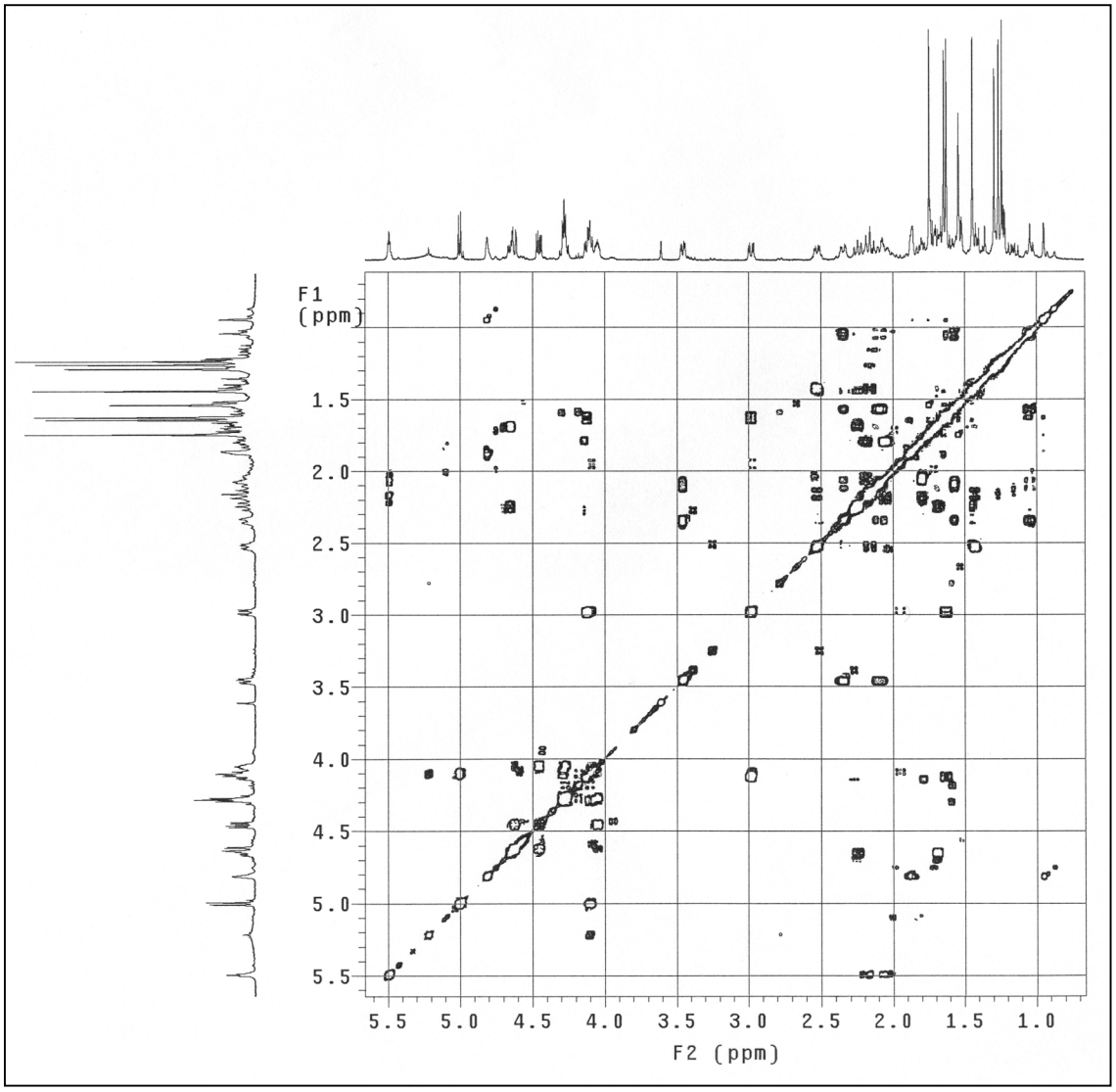


Figure A.19. COSY spectra for *S. radula* compound **4**.

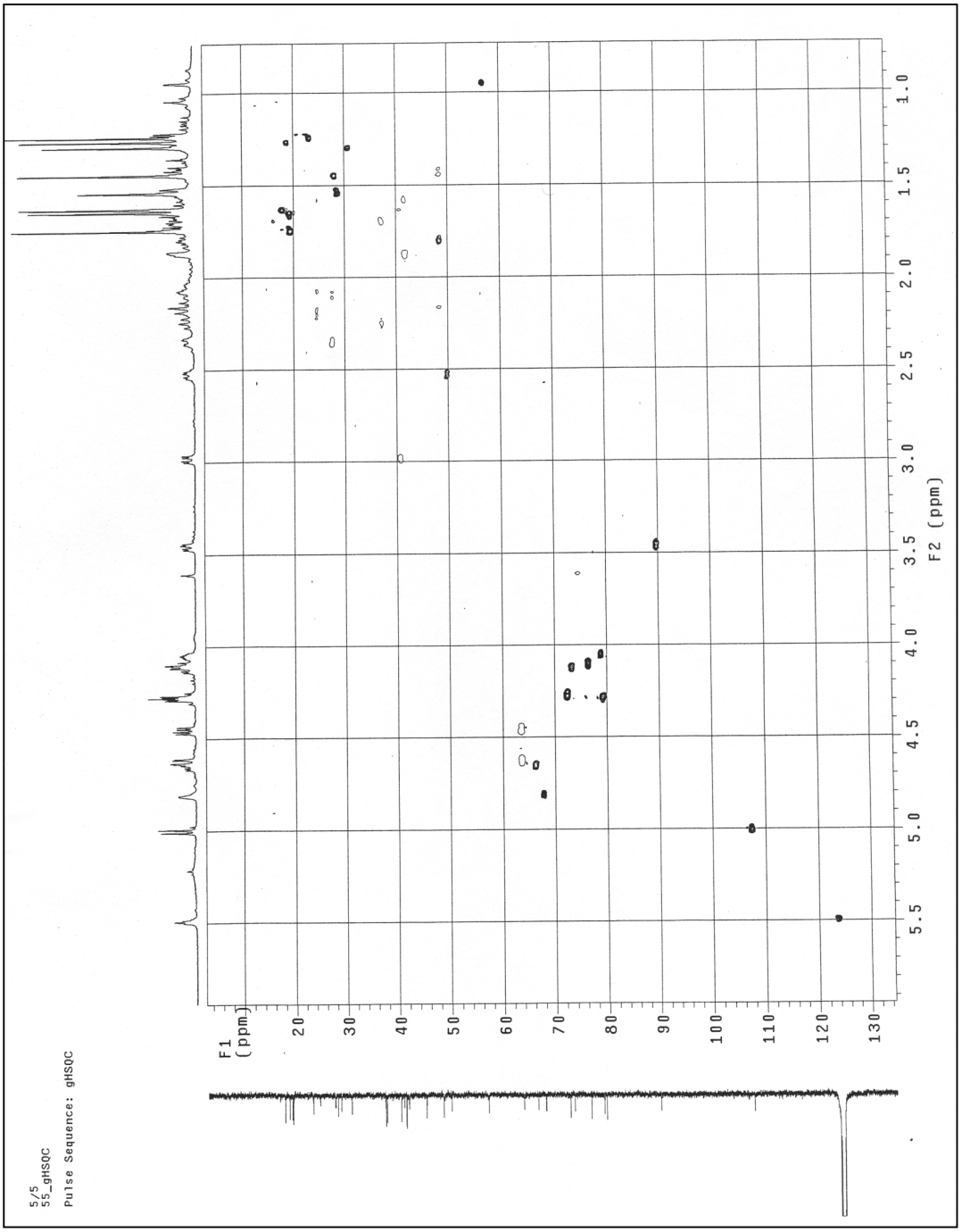


Figure A.20. HSQC spectra for *S. radula* compound **4**.

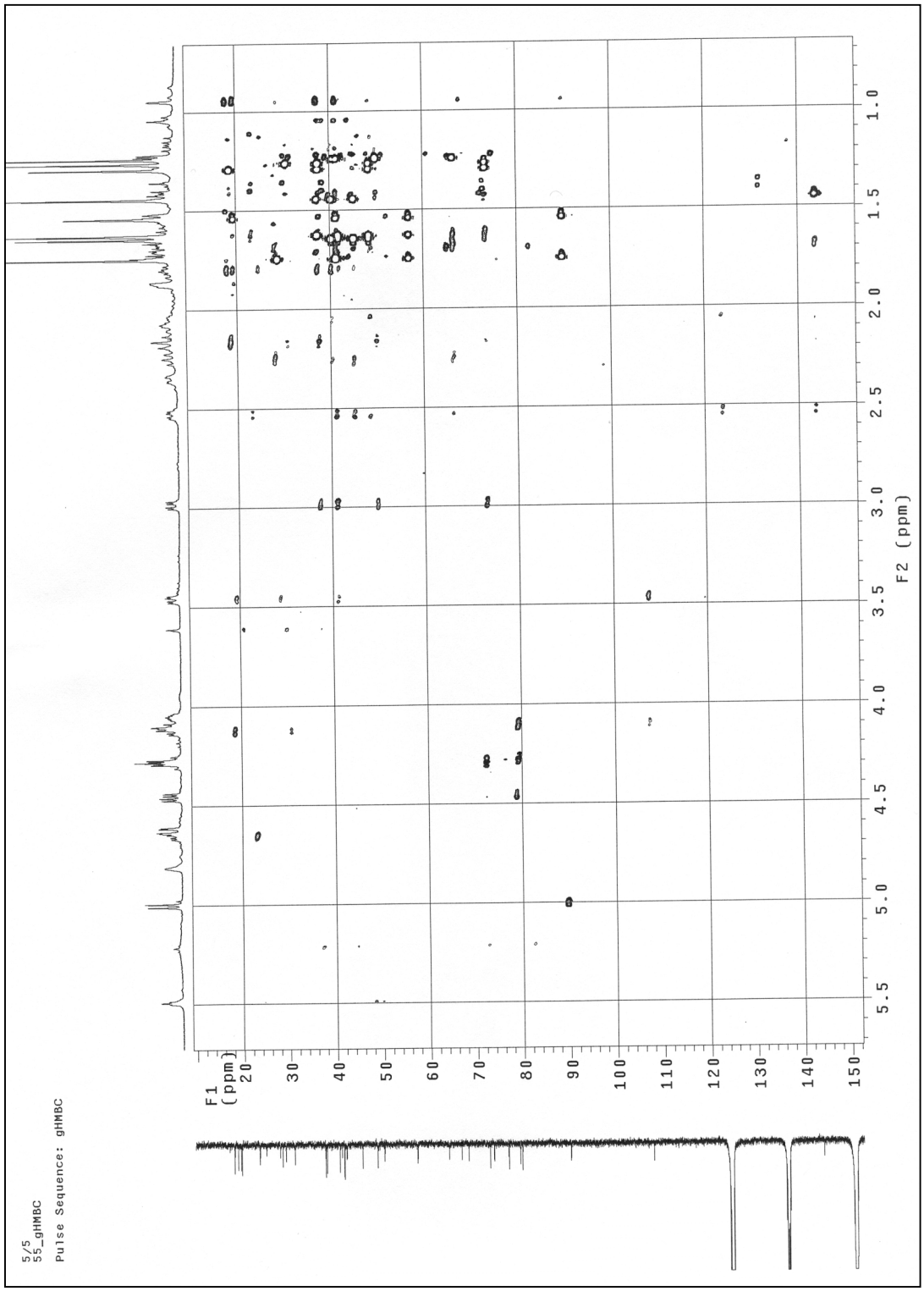


Figure A.21. HMBC spectra for *S. radula* compound **4**.

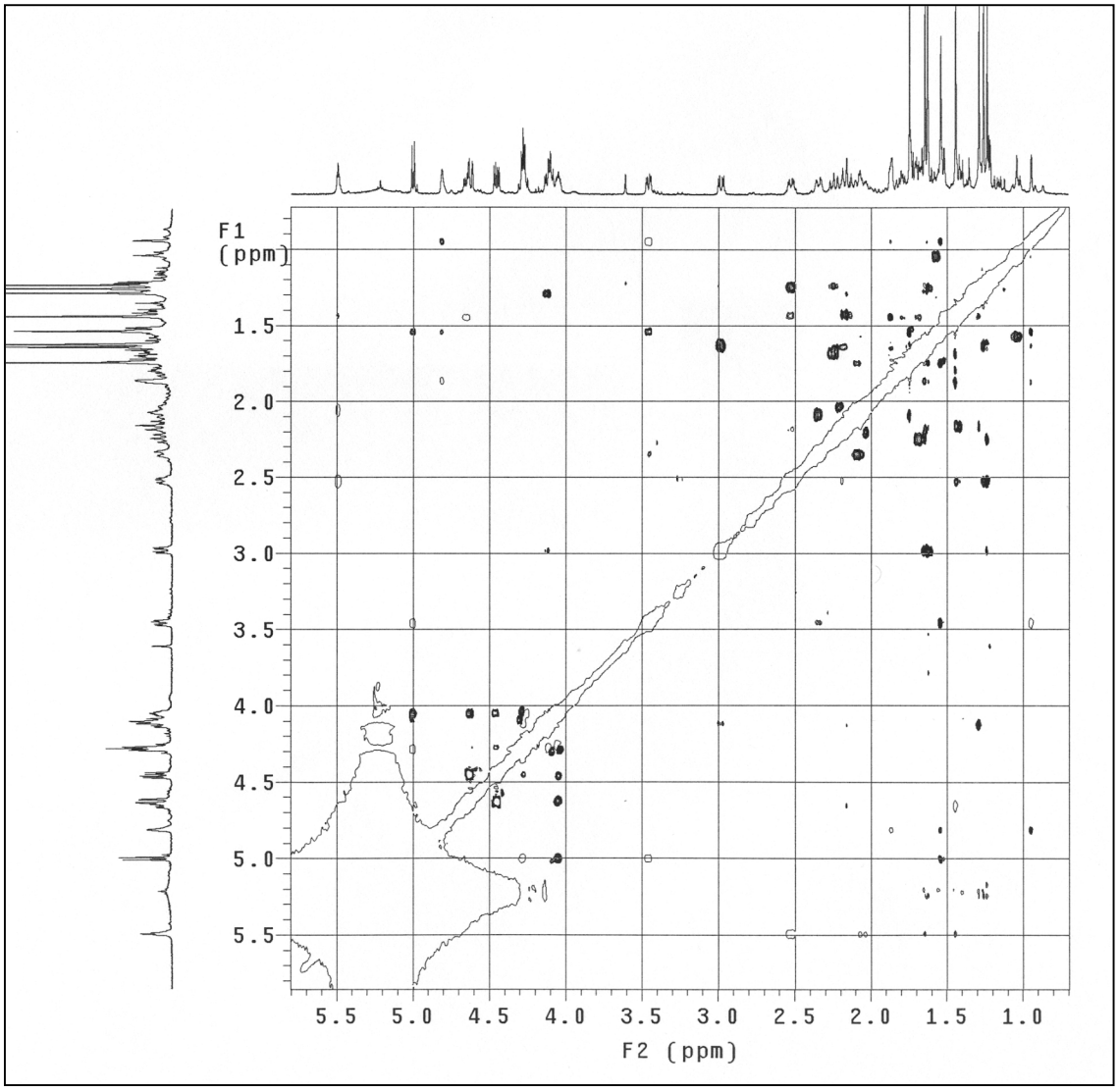


Figure A.22. NOESY spectra for *S. radula* compound **4**.

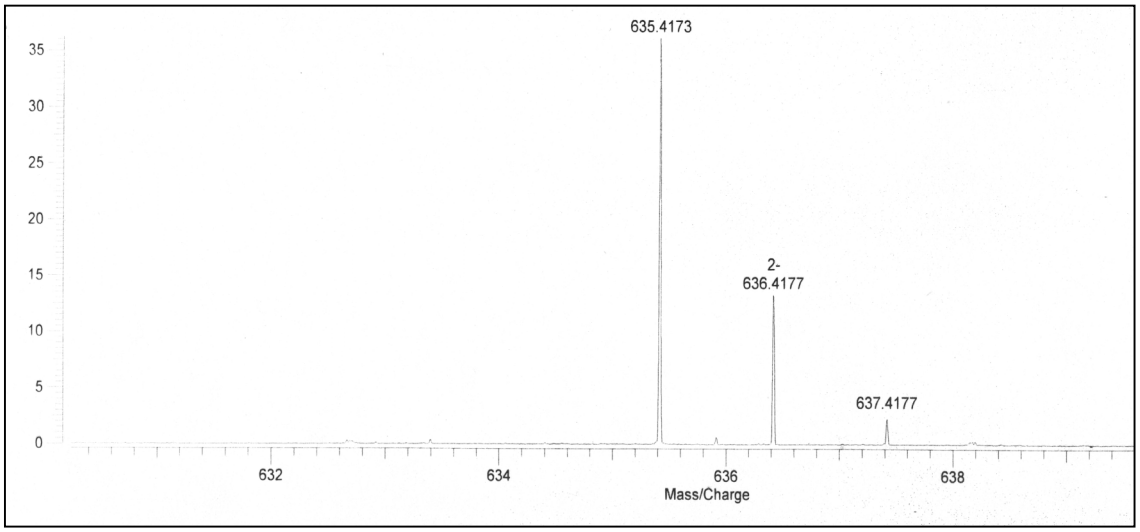


Figure A.23. HR-ESI-MS negative mode spectra for *S. radula* compound **5**.

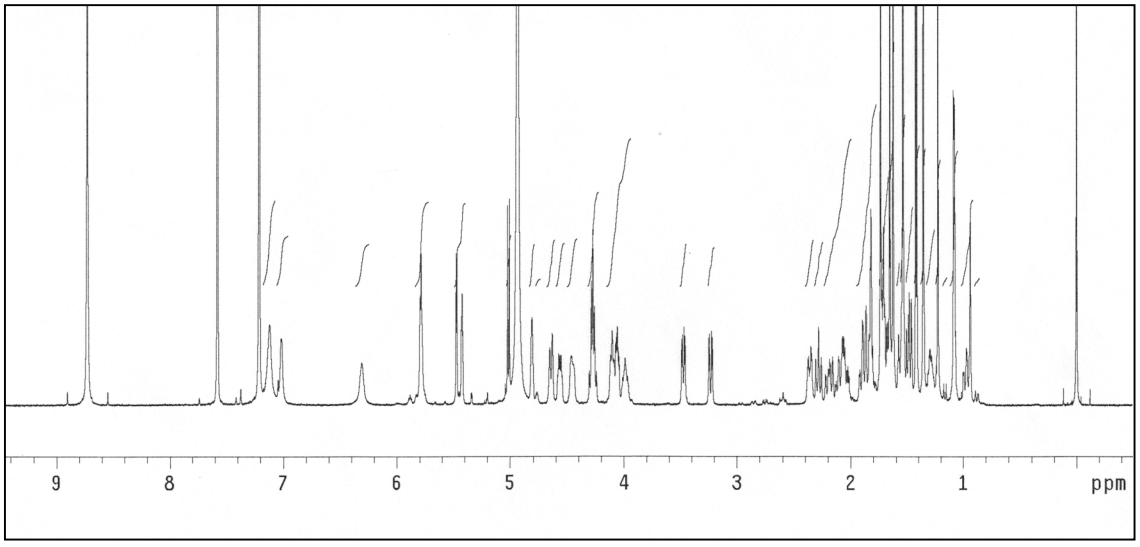


Figure A.24. ^{1}H-NMR spectra for *S. radula* compound **5**.

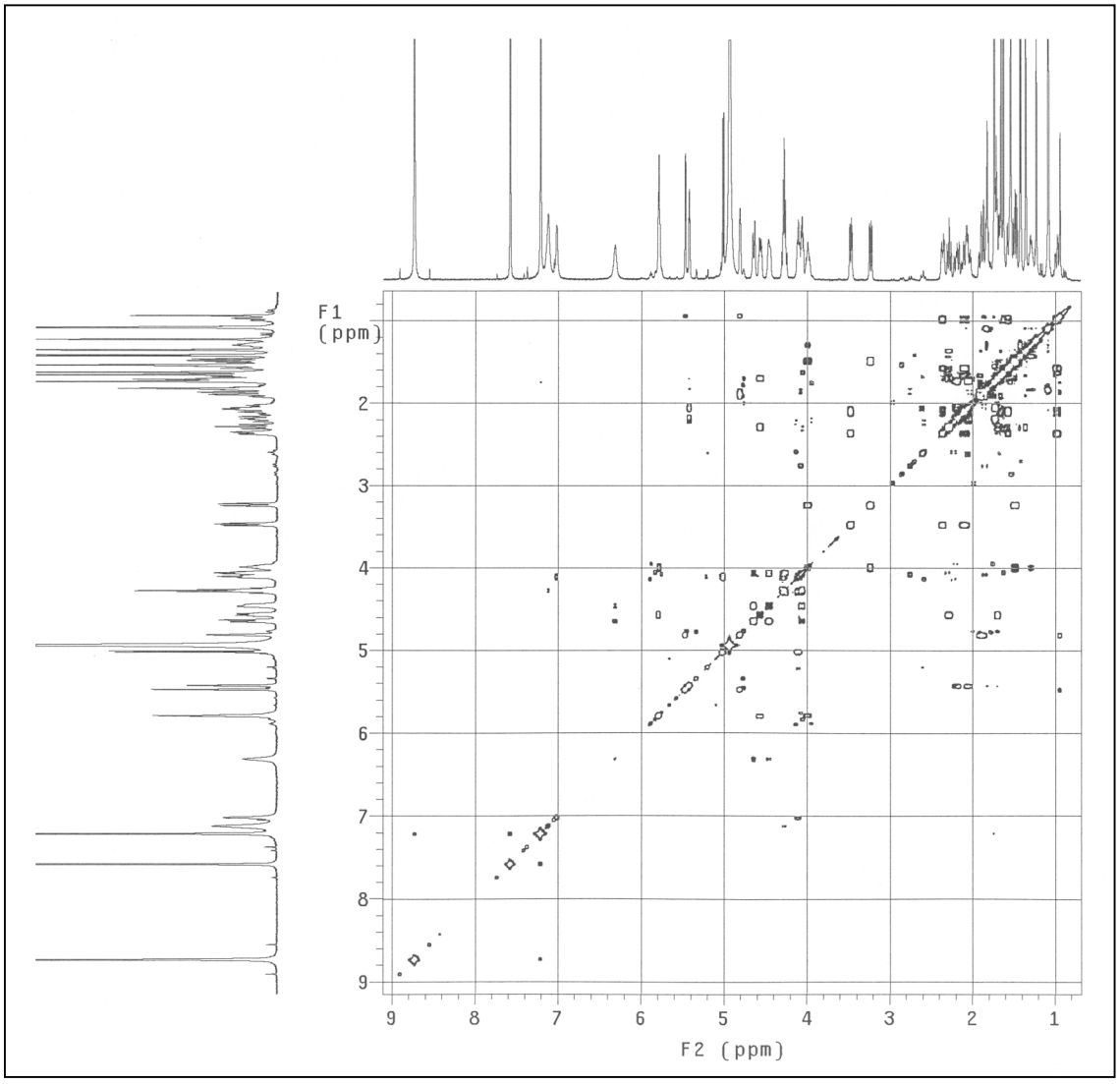


Figure A.25. COSY spectra for *S. radula* compound **5**.

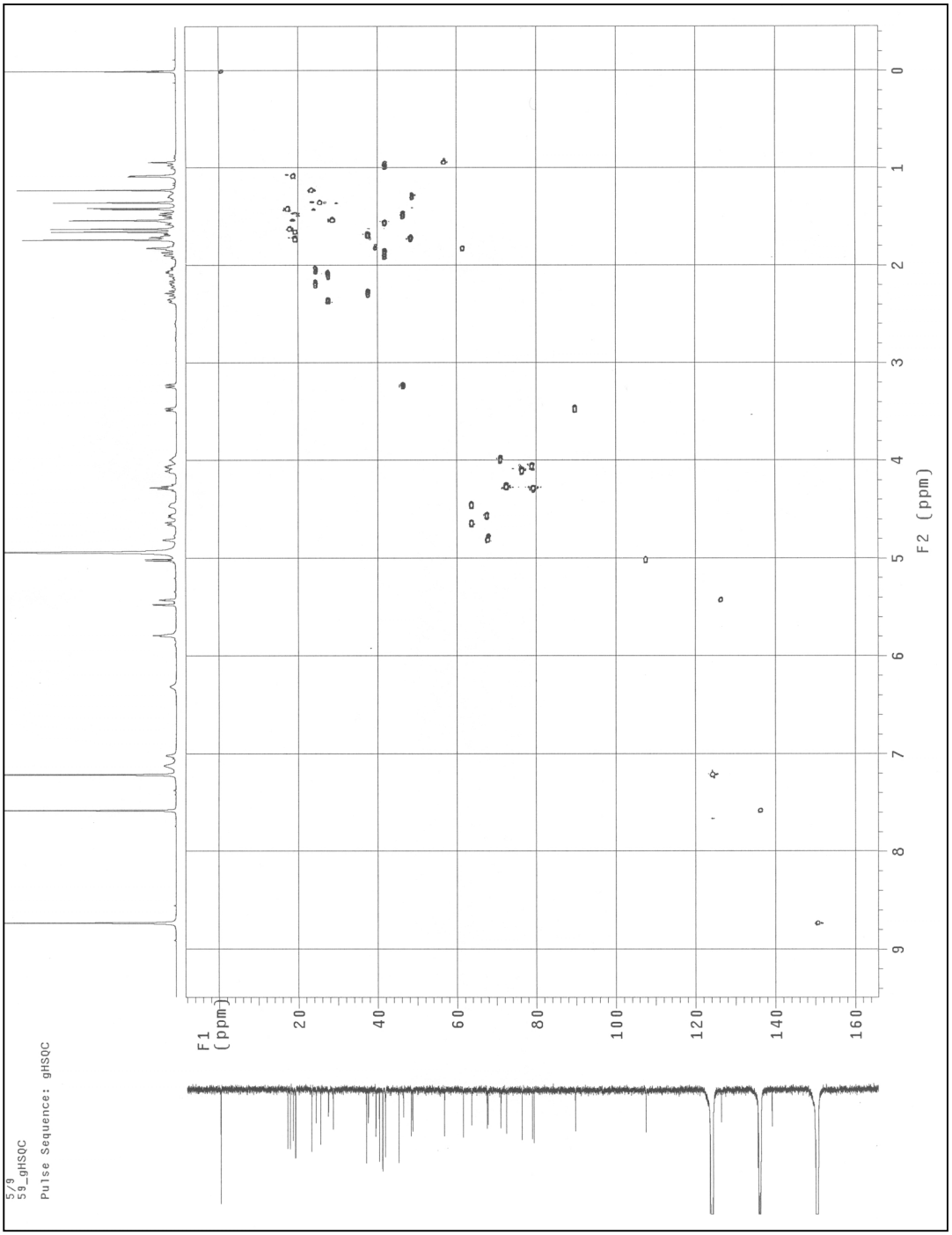


Figure A.26. HSQC spectra for *S. radula* compound **5**.

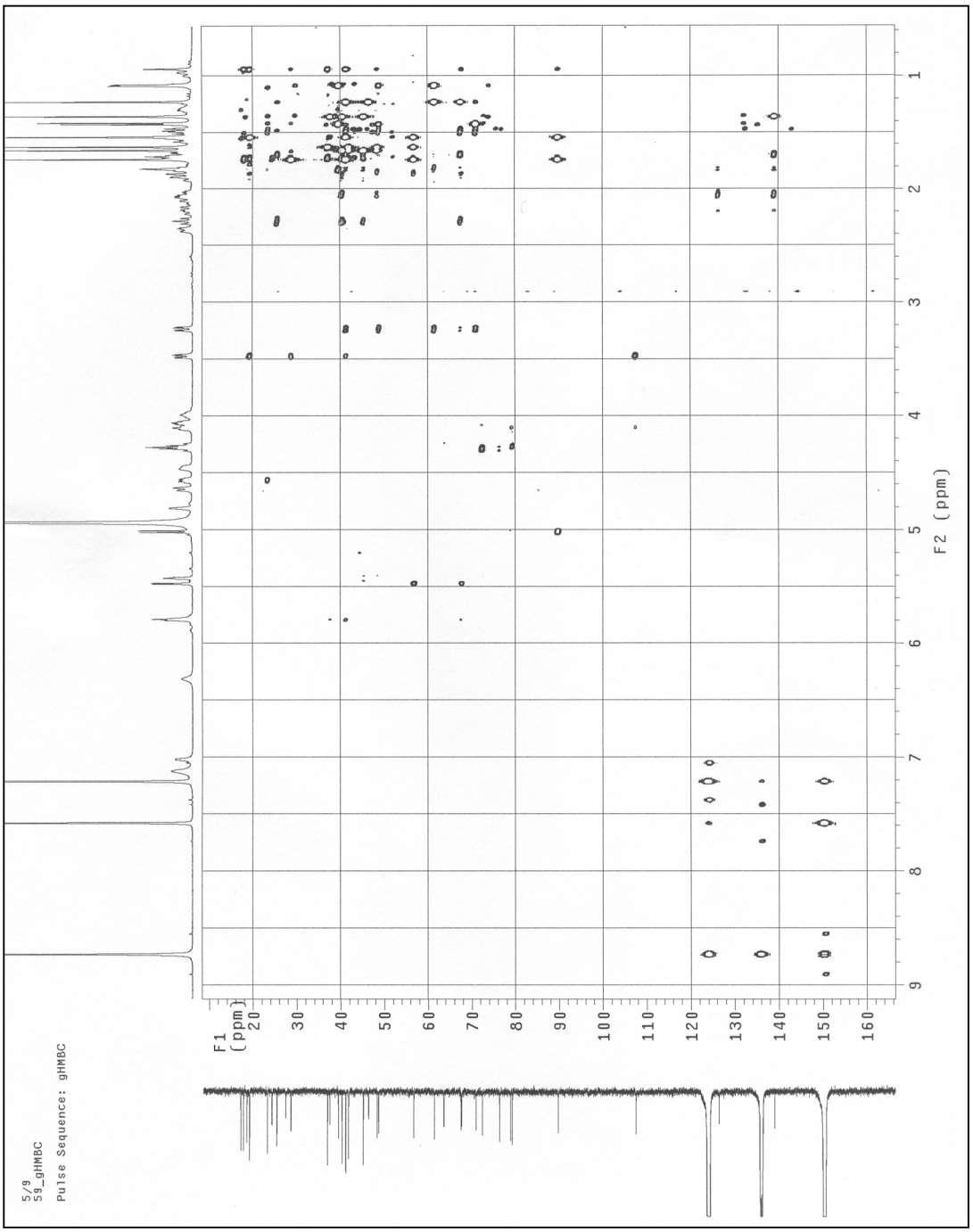


Figure A.27. HMBC spectra for *S. radula* compound **5**.

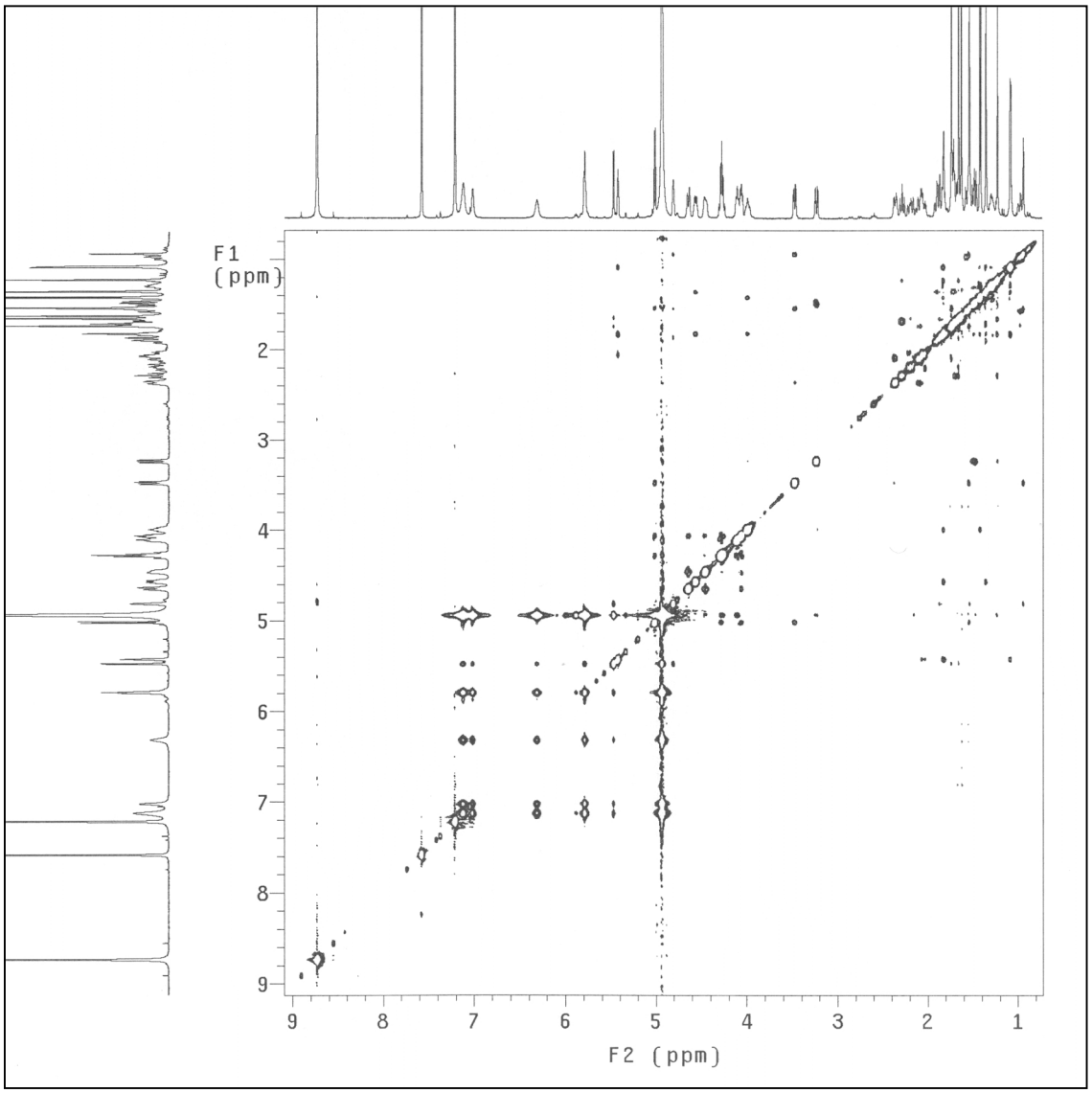


Figure A.28. NOESY spectra for *S. radula* compound **5**.

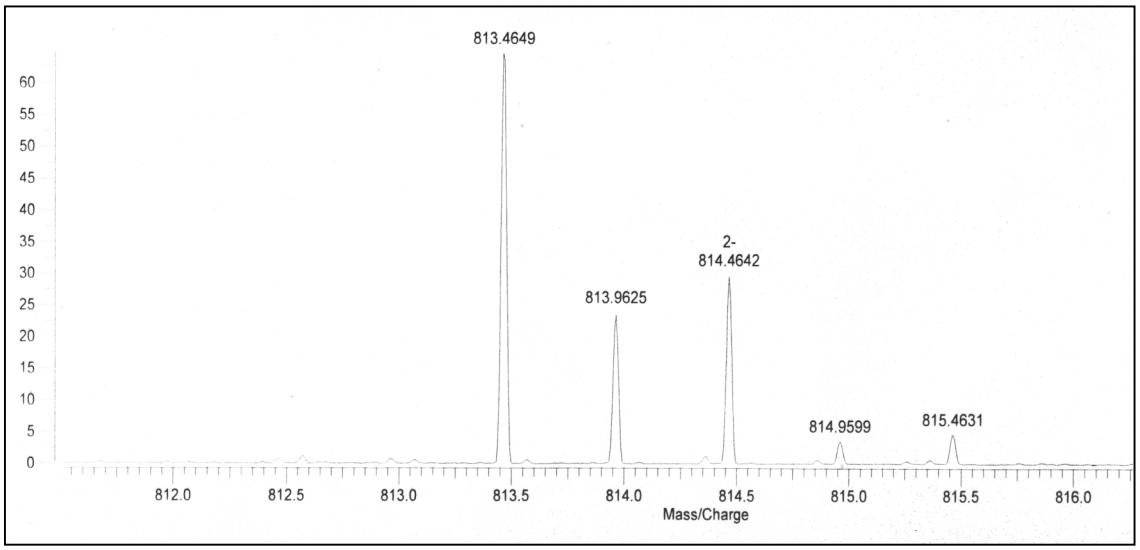


Figure A.29. HR-ESI-MS negative mode spectra for *S. radula* compound **6**.

.

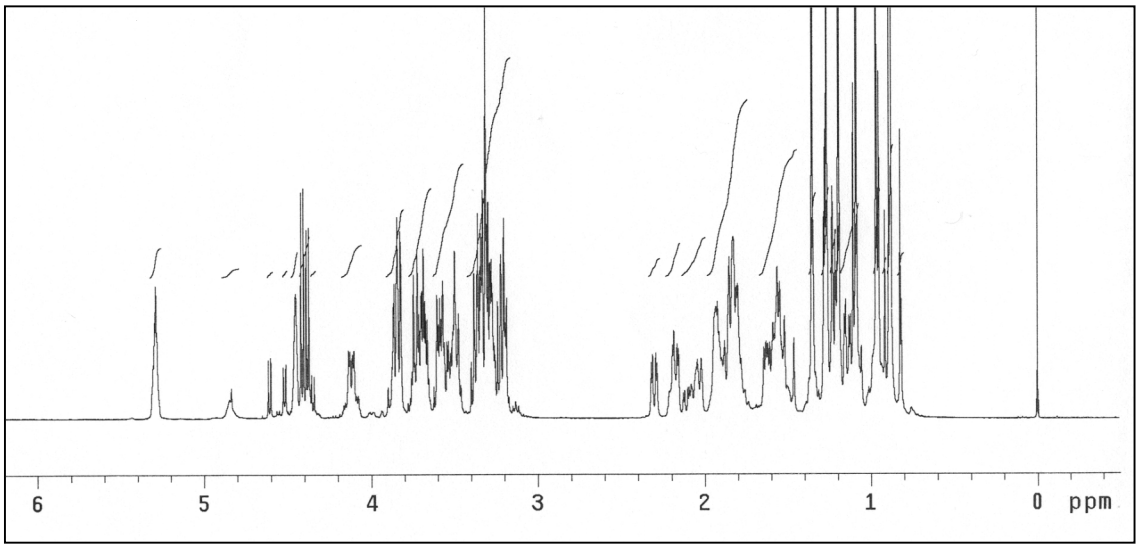


Figure A.30. ^{1}H-NMR spectra for *S. radula* compound **6.**

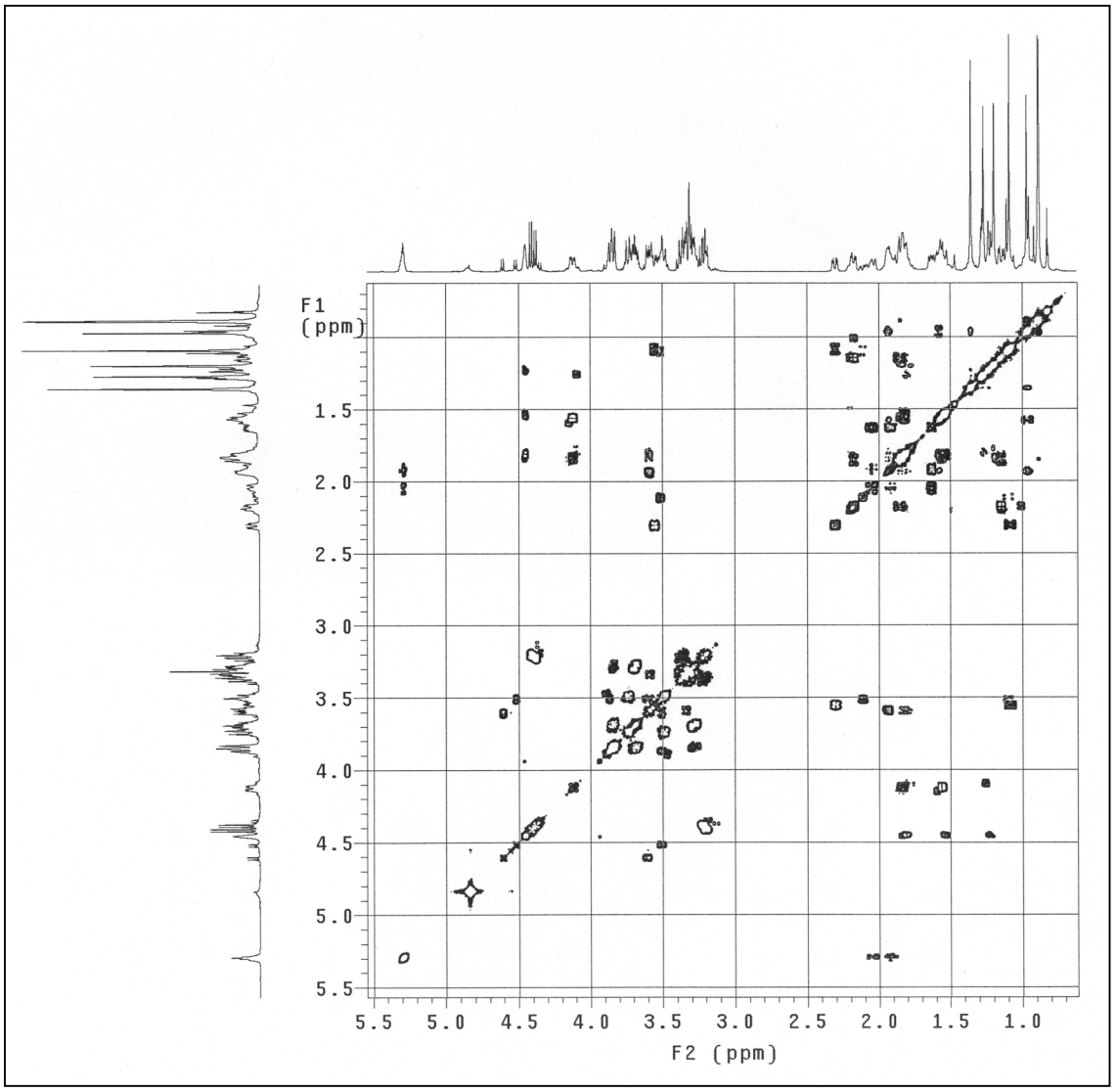


Figure A.31. COSY spectra for *S. radula* compound **6**.

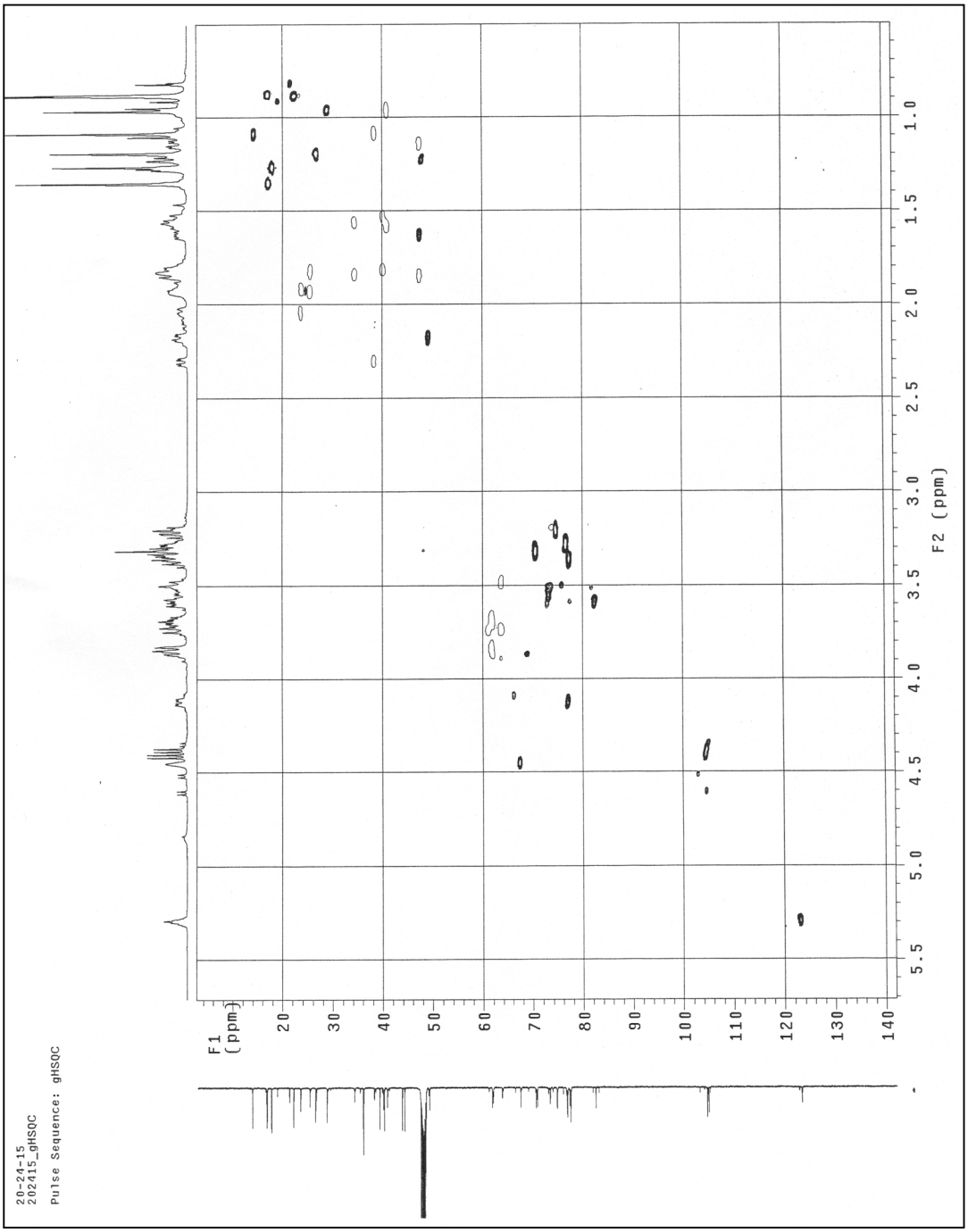


Figure A.32. HSQC spectra for *S. radula* compound **6**.

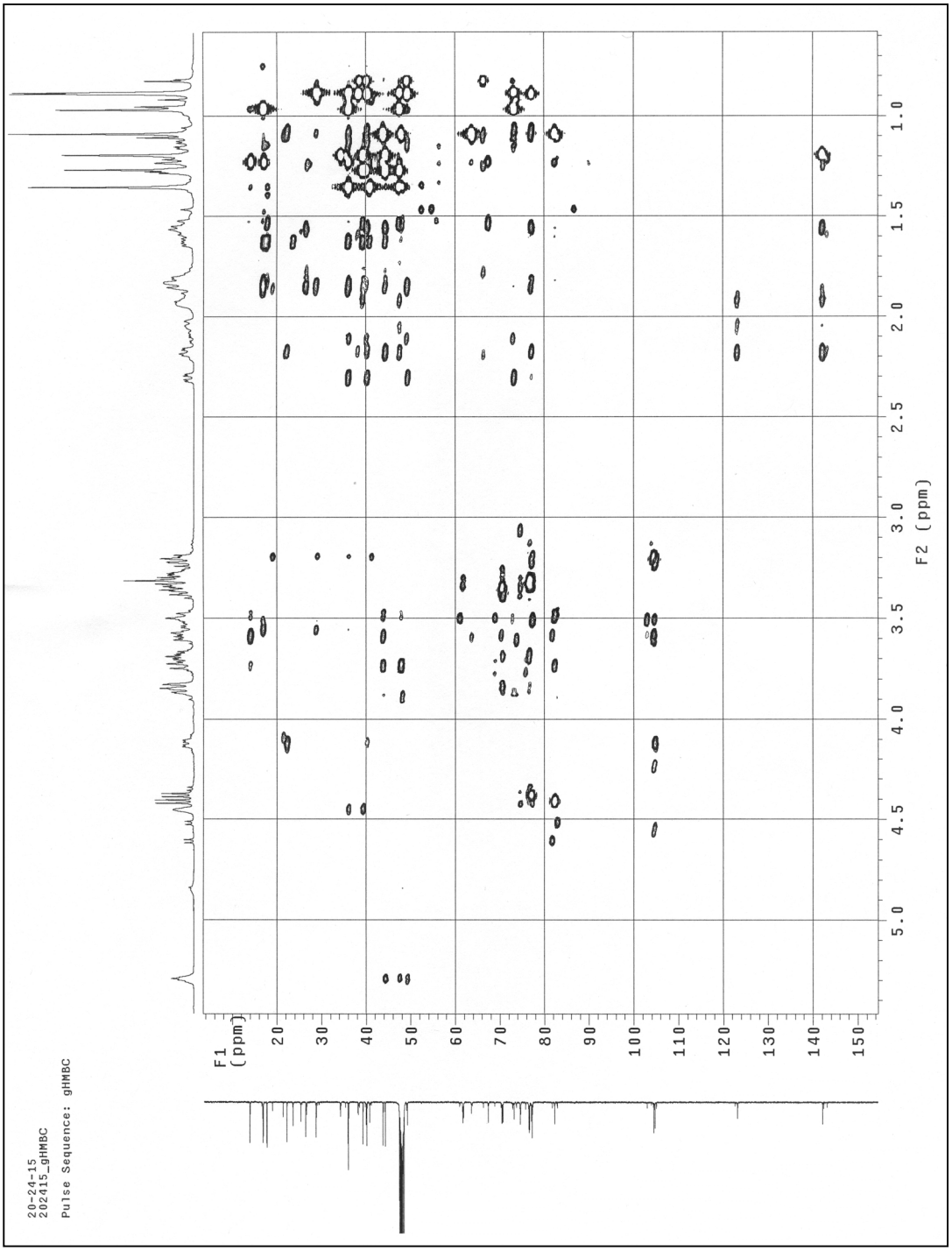


Figure A.33. HMBC spectra for *S. radula* compound **6.**

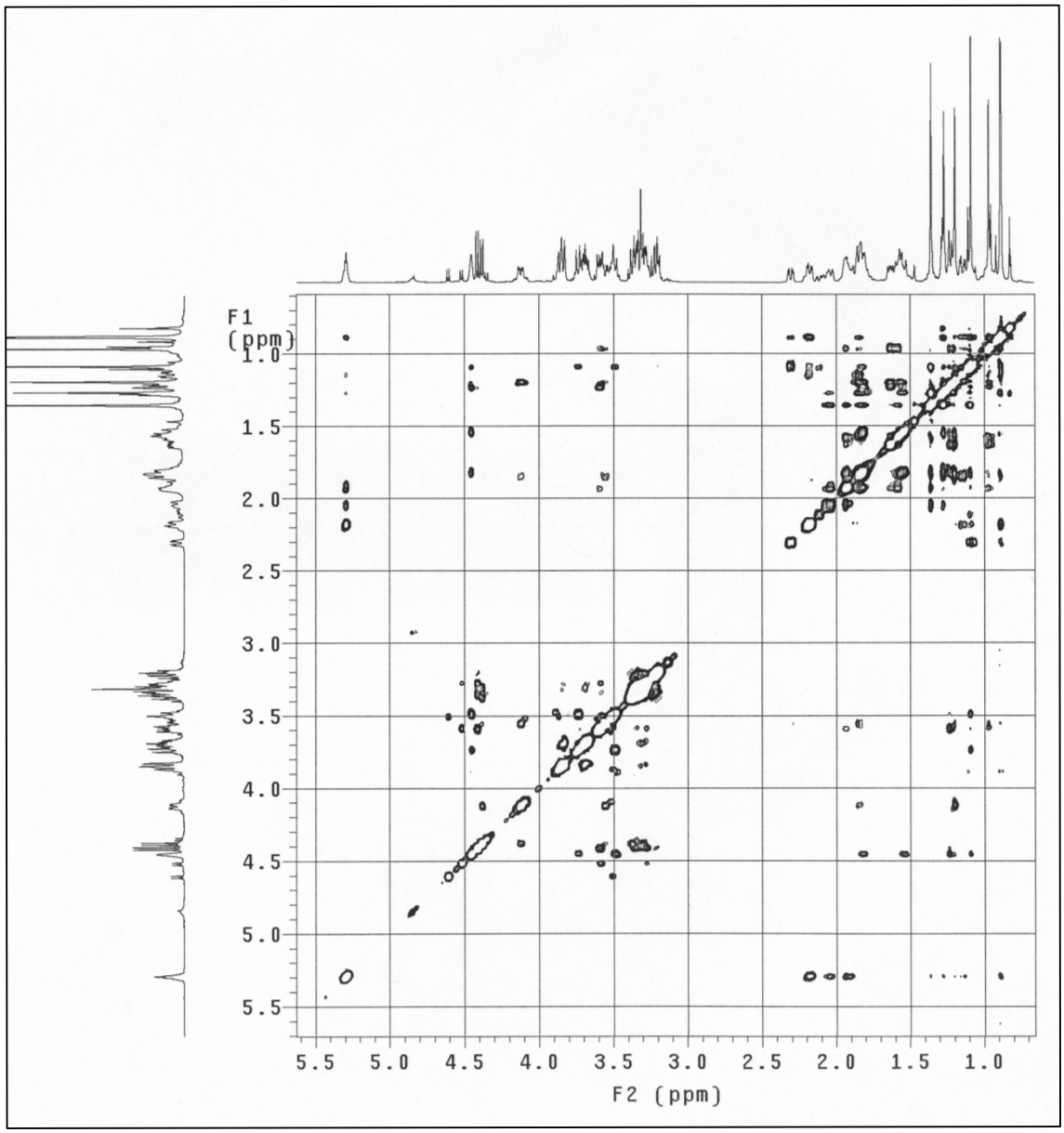


Figure A.34. NOESY spectra for *S. radula* compound **6.**

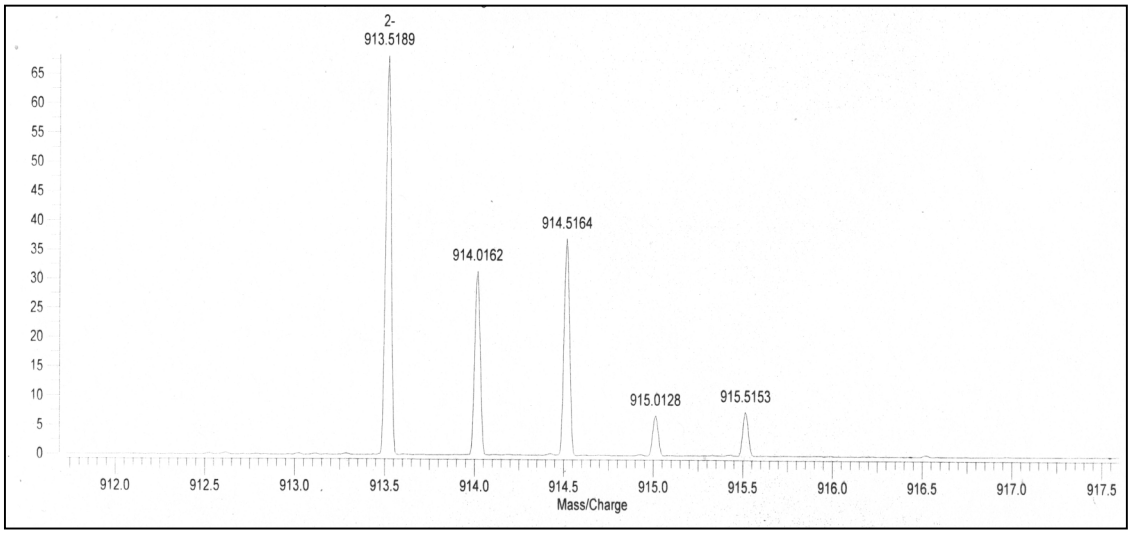


Figure A.35. HR-ESI-MS negative mode spectra for *S. radula* compound **7.**

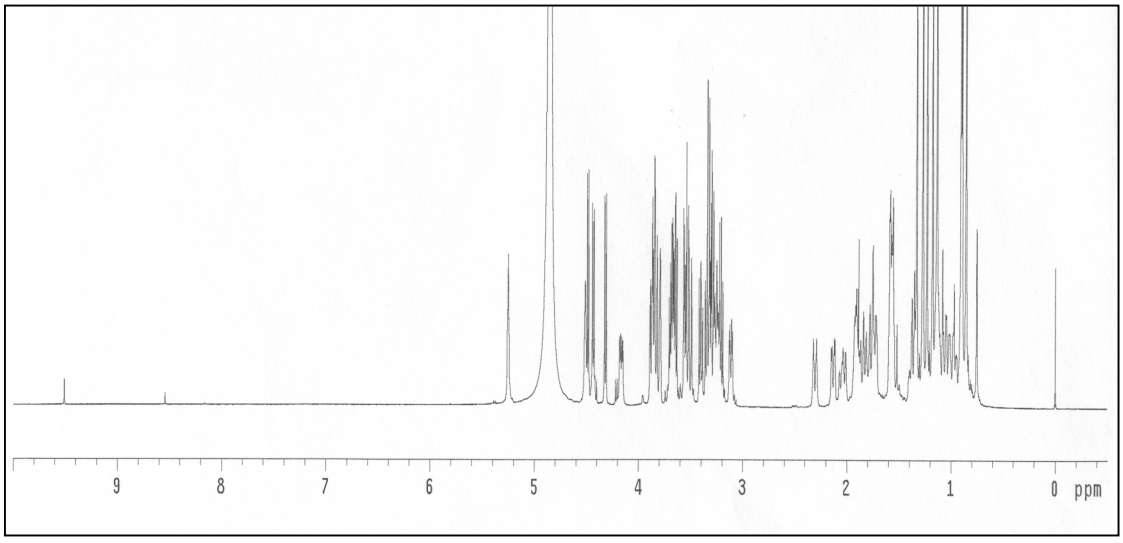


Figure A.36. ^{1}H-NMR spectra for *S. radula* compound **7.**

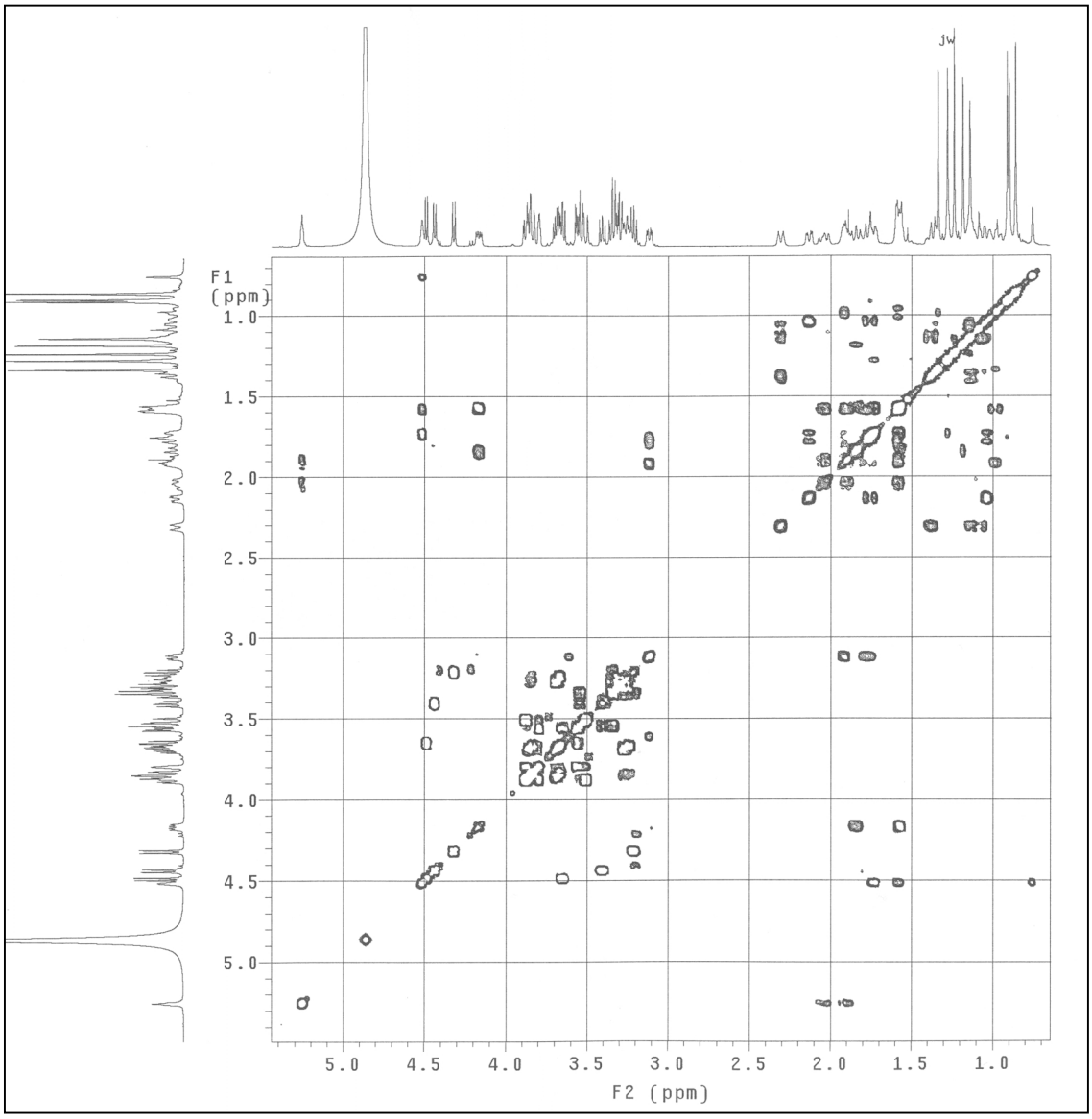


Figure A.37. COSY spectra for *S. radula* compound **7.**

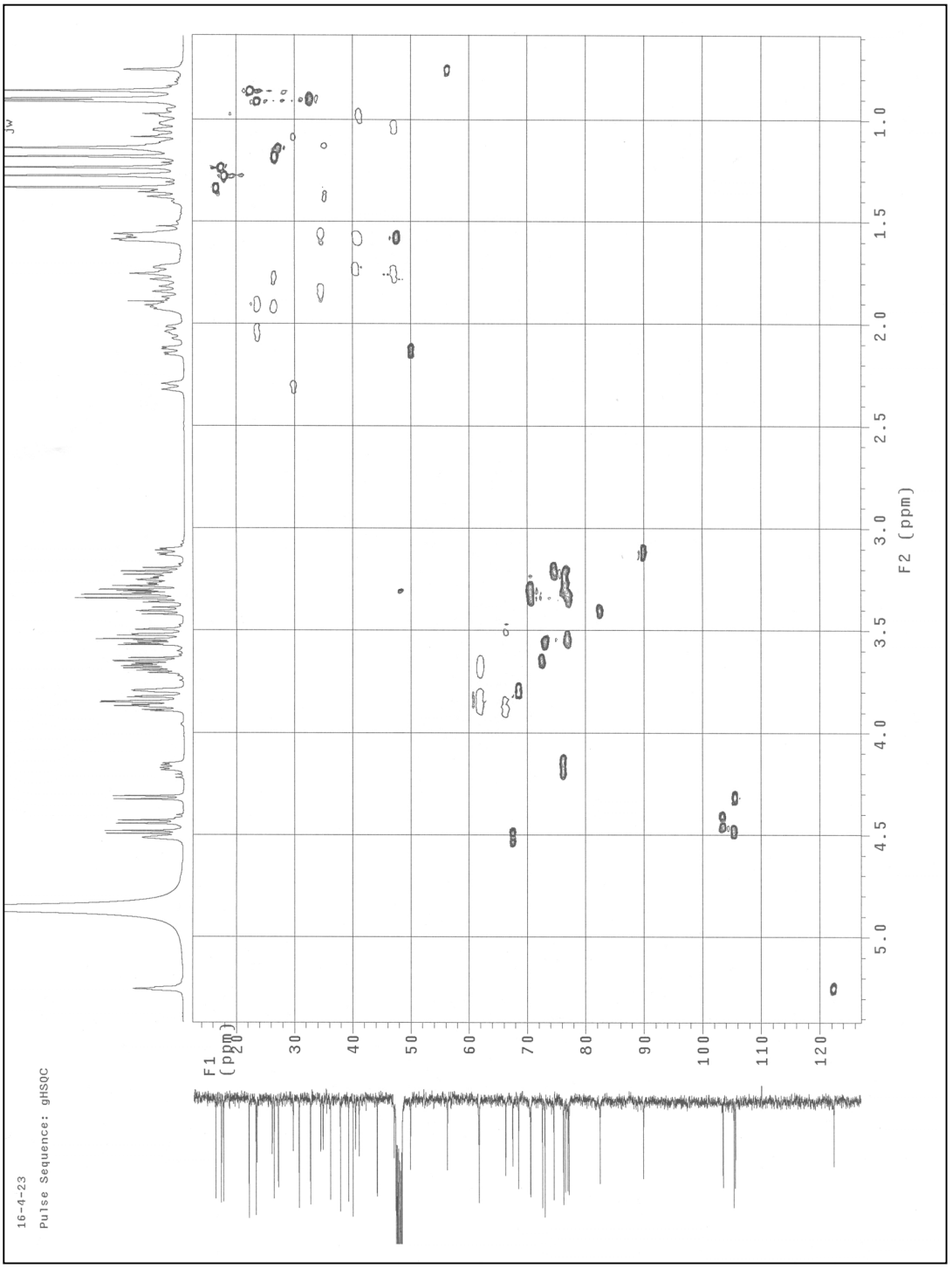


Figure A.38. HSQC spectra for *S. radula* compound **7.**

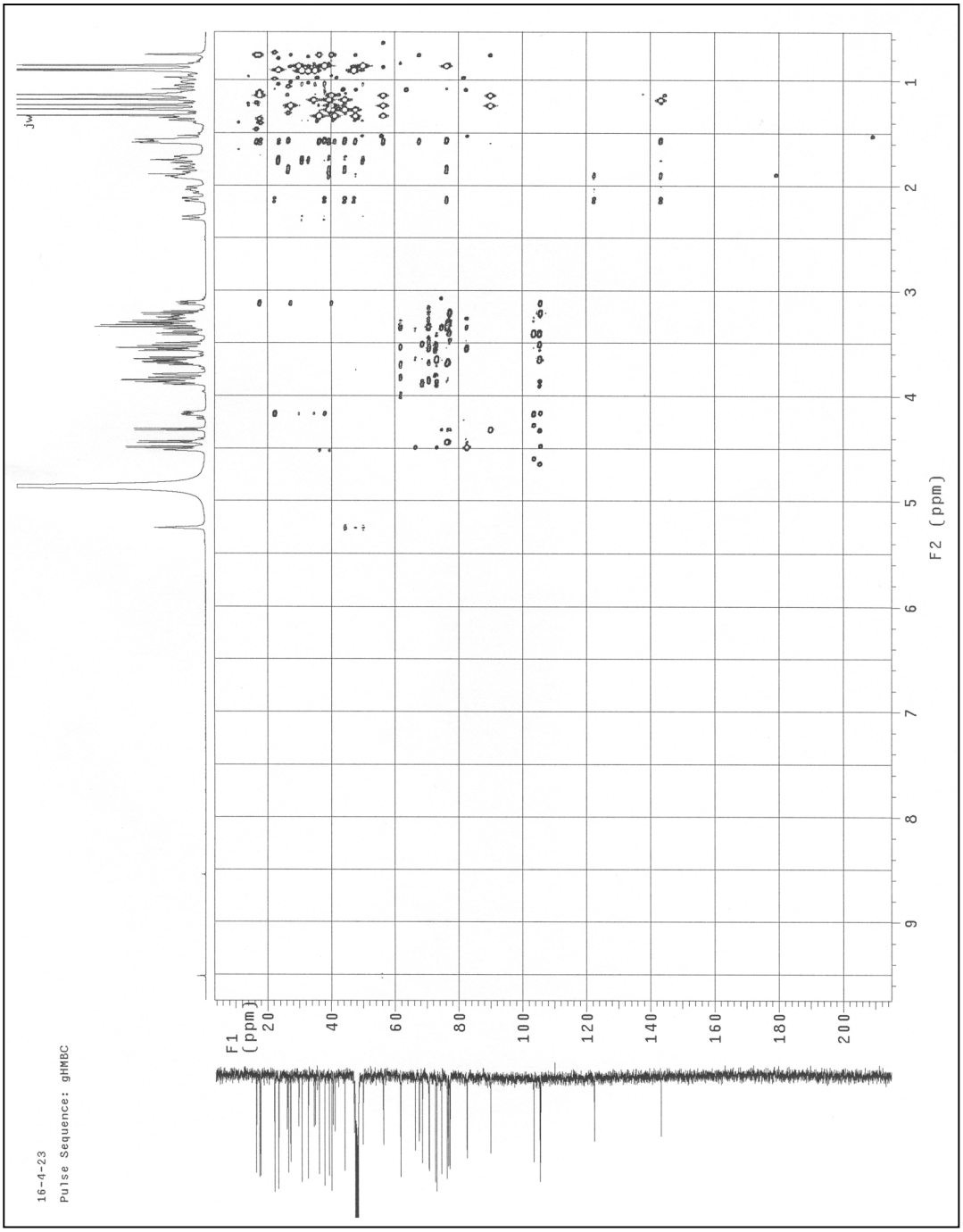


Figure A.39. HMBC spectra for *S. radula* compound **7**.

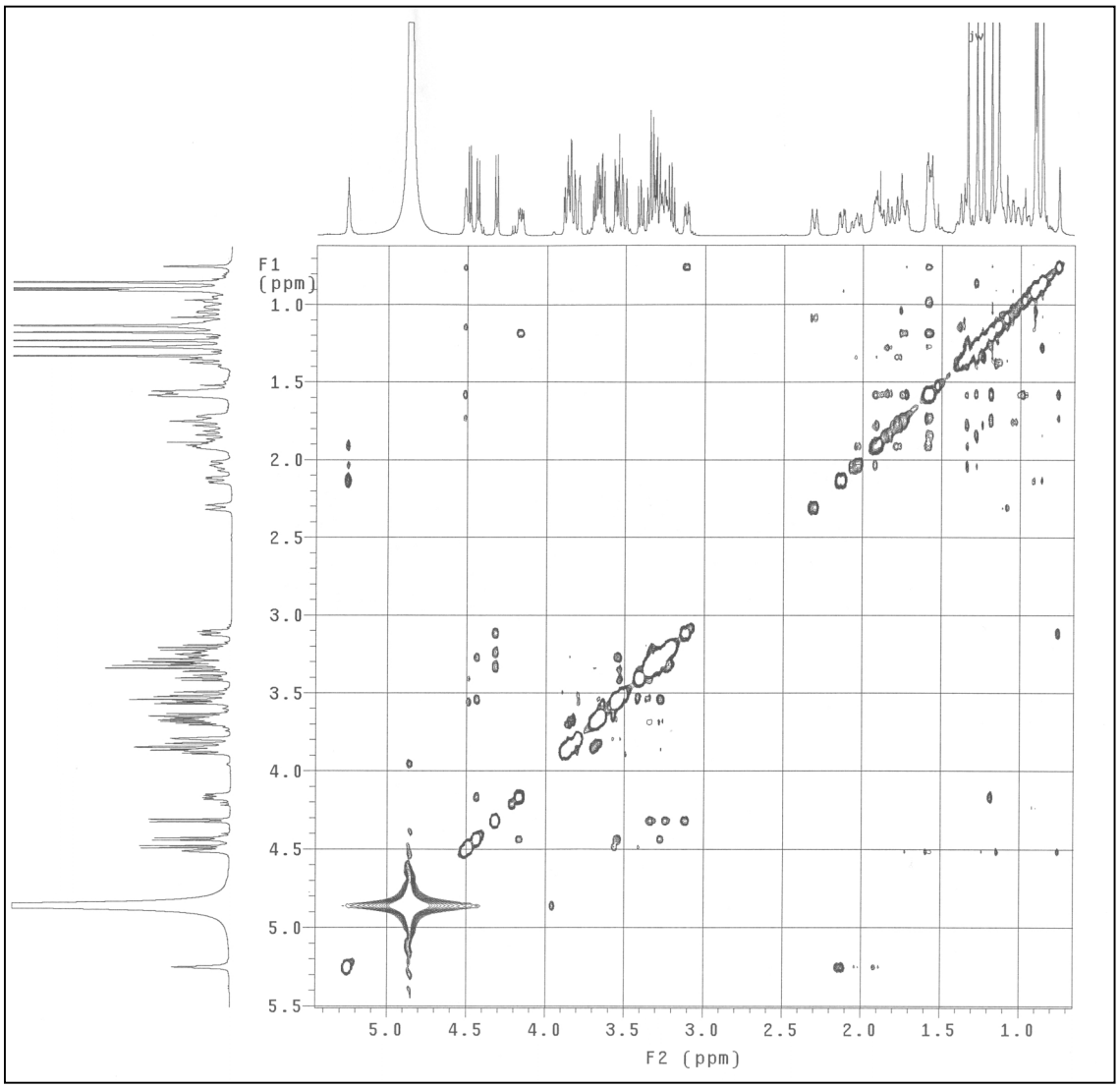
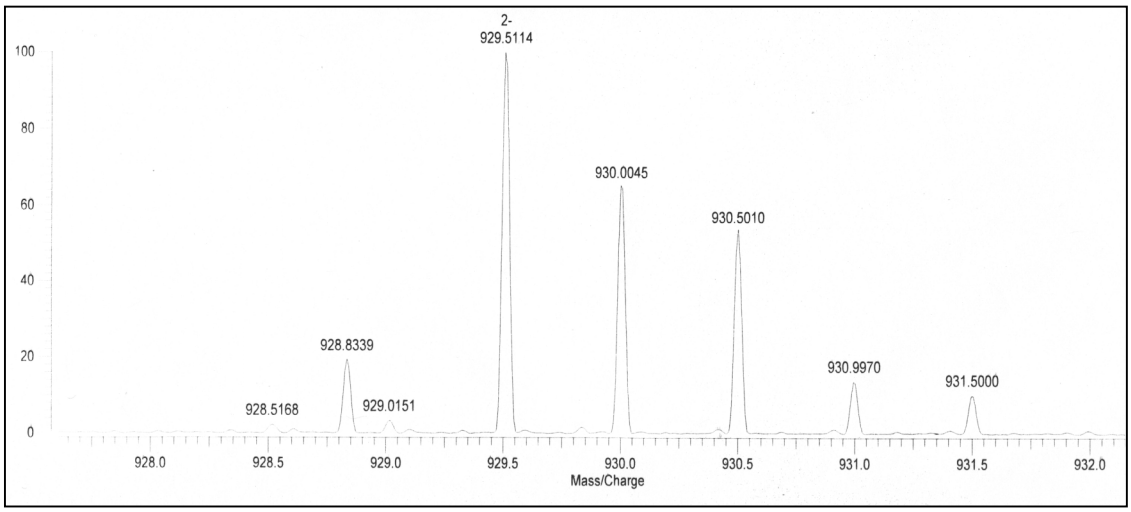
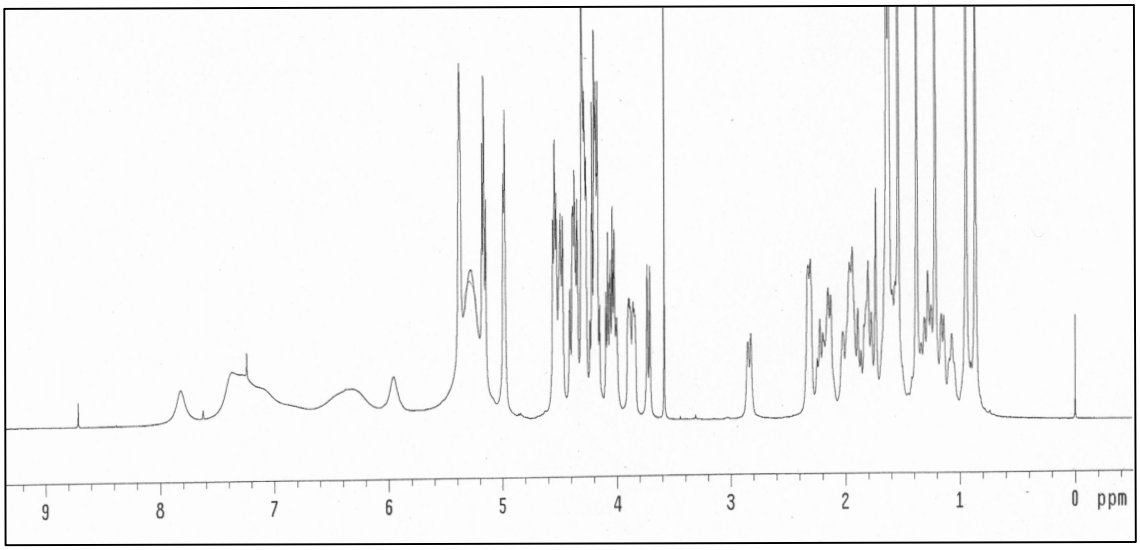


Figure A.40. NOESY spectra for *S. radula* compound **7**.

.



A.41. HR-ESI-MS negative mode spectra for *S. radula* compound **8.**



A.42. ^{1}H-NMR spectra for *S. radula* compound **8.**

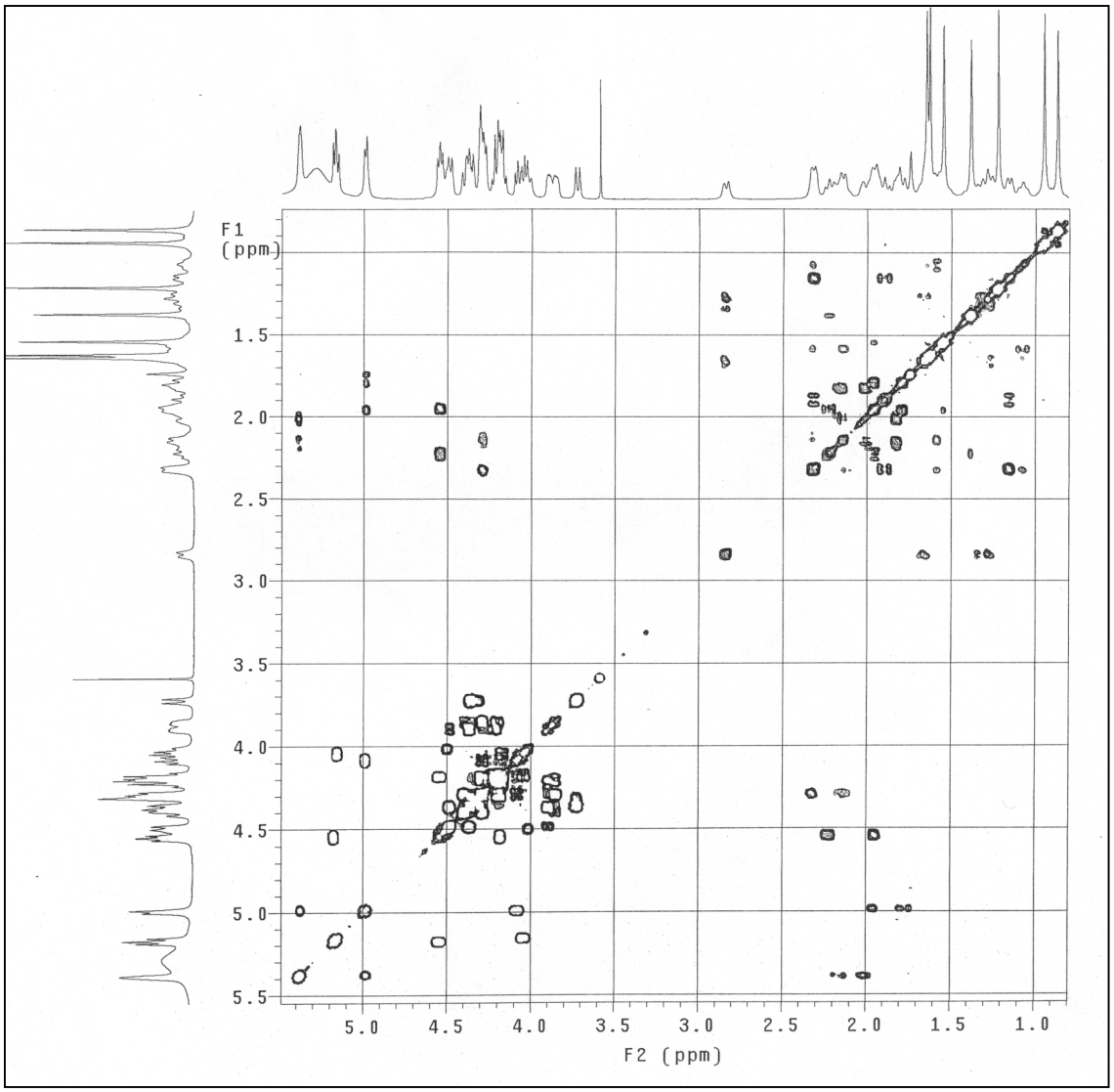


Figure A.43. COSY spectra for *S. radula* compound **8.**

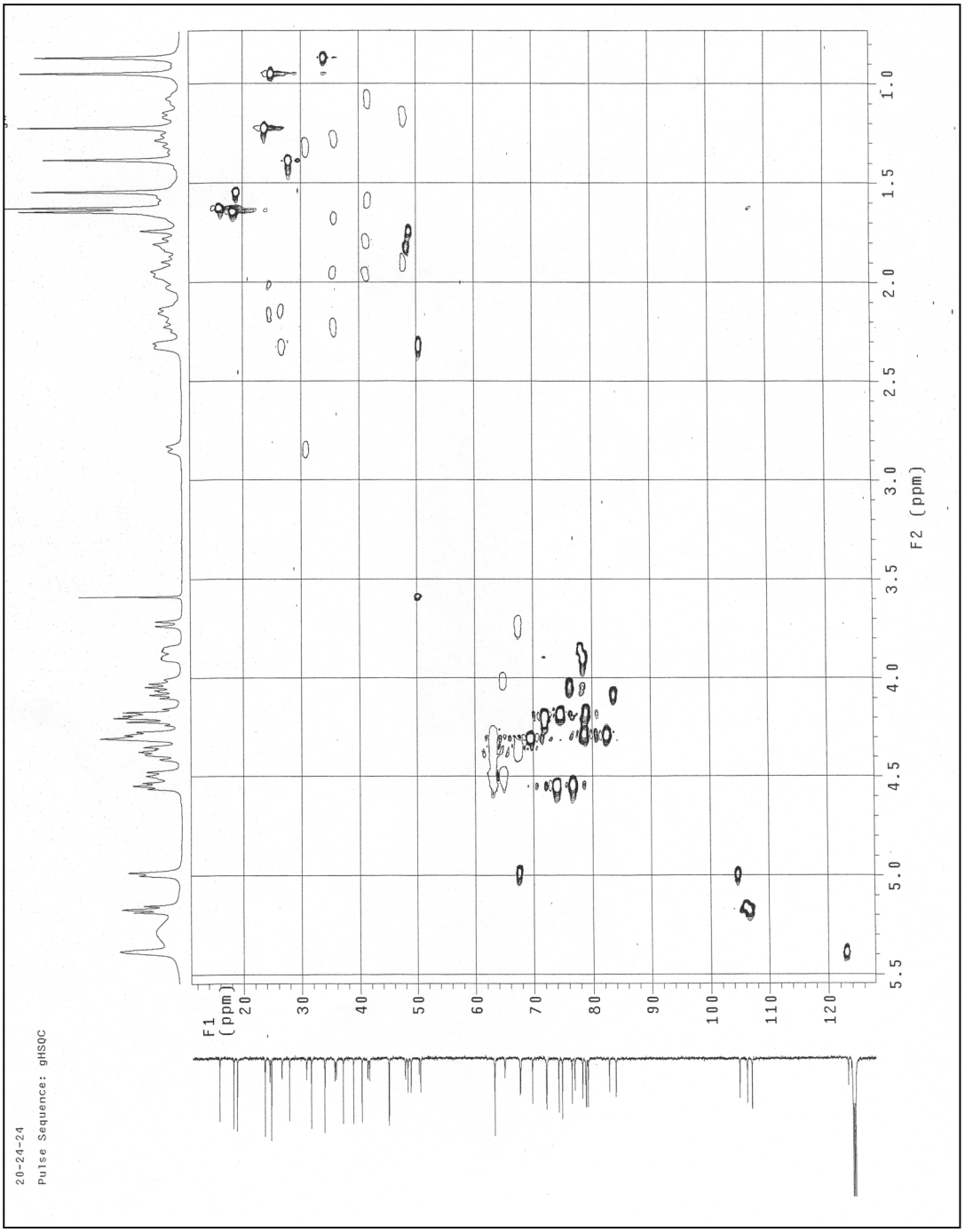


Figure A.44. HSQC spectra for *S. radula* compound **8**.

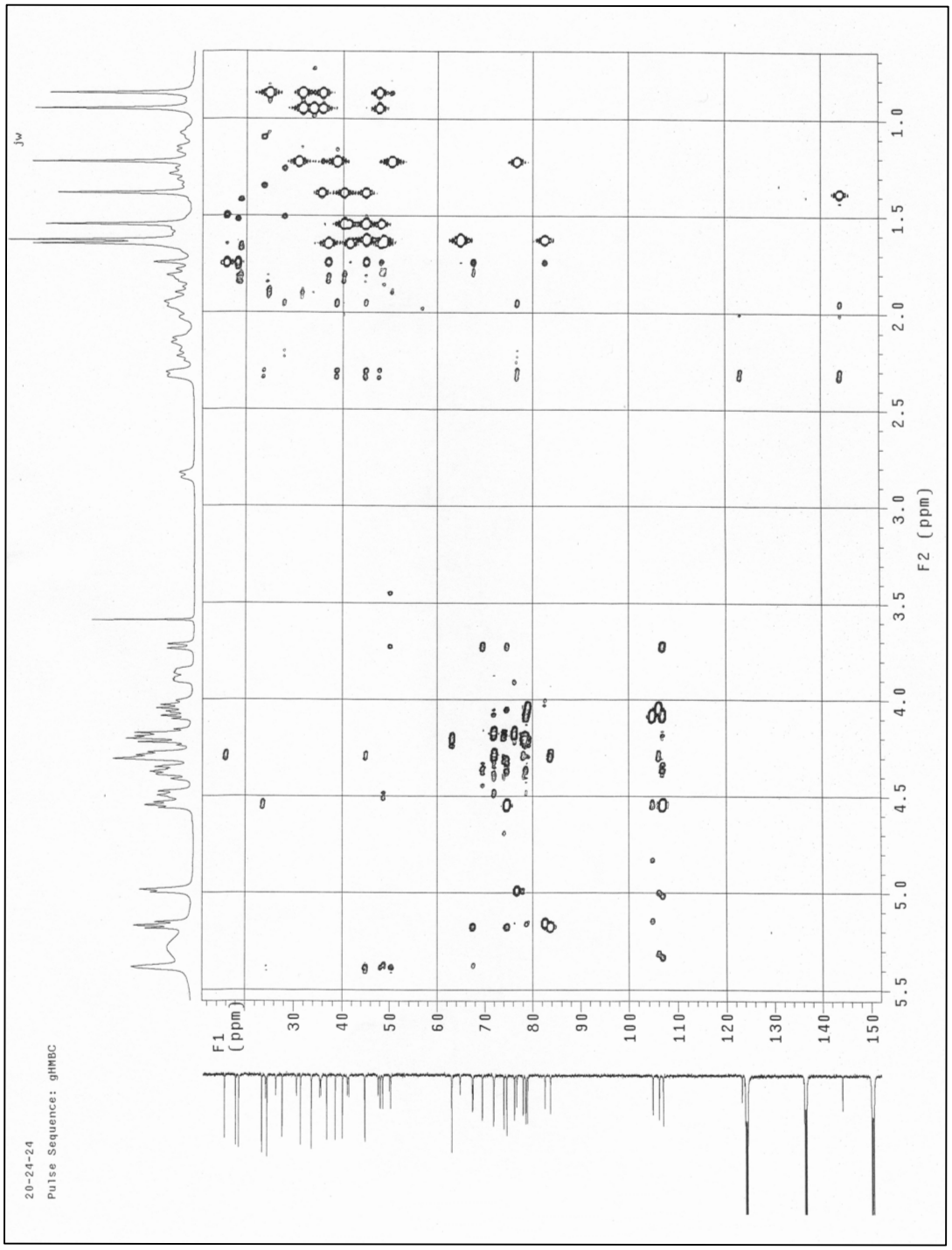


Figure A.45. HMBC spectra for *S. radula* compound **8.**

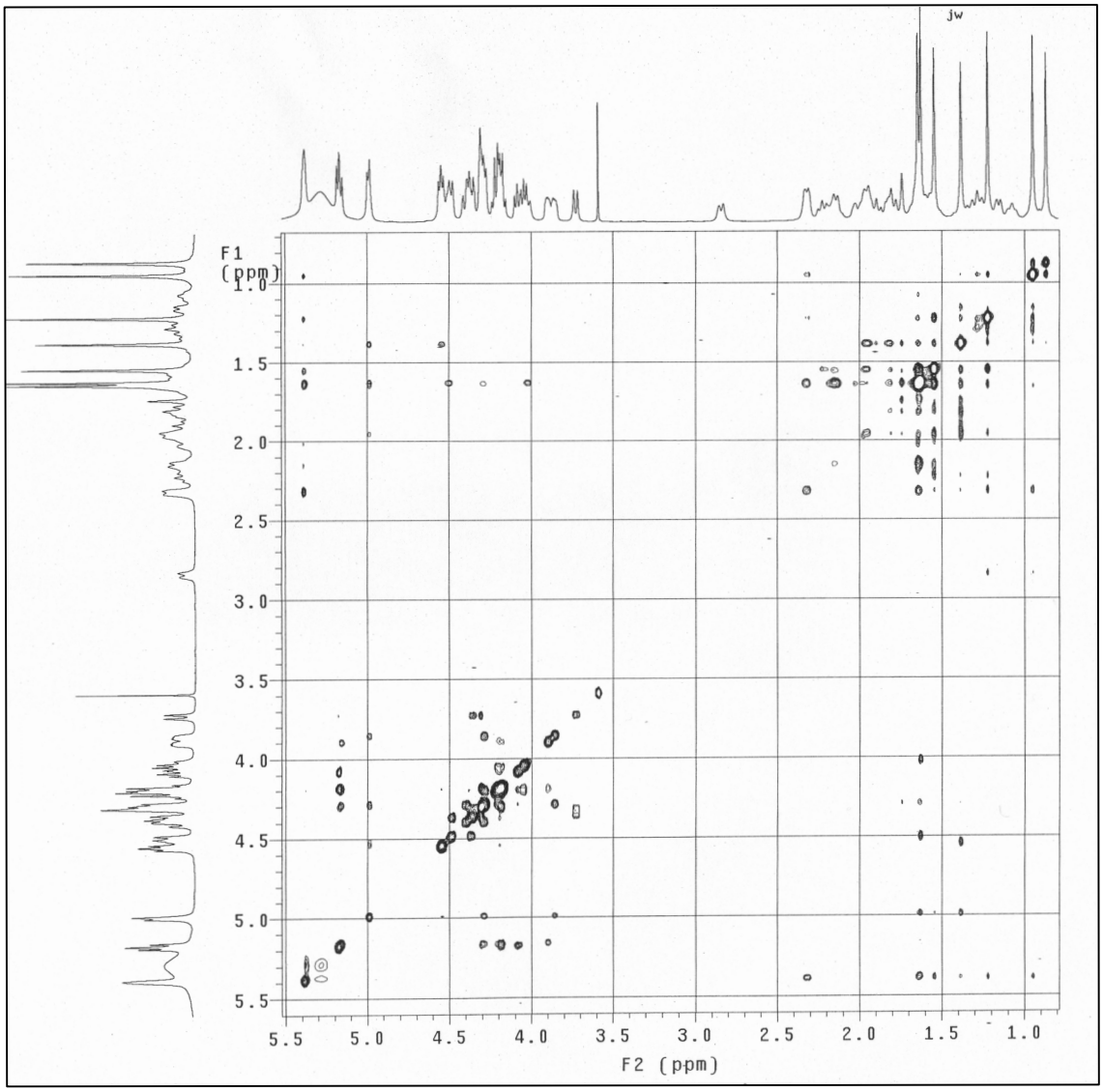


Figure A.46. NOESY spectra for *S. radula* compound **8.**

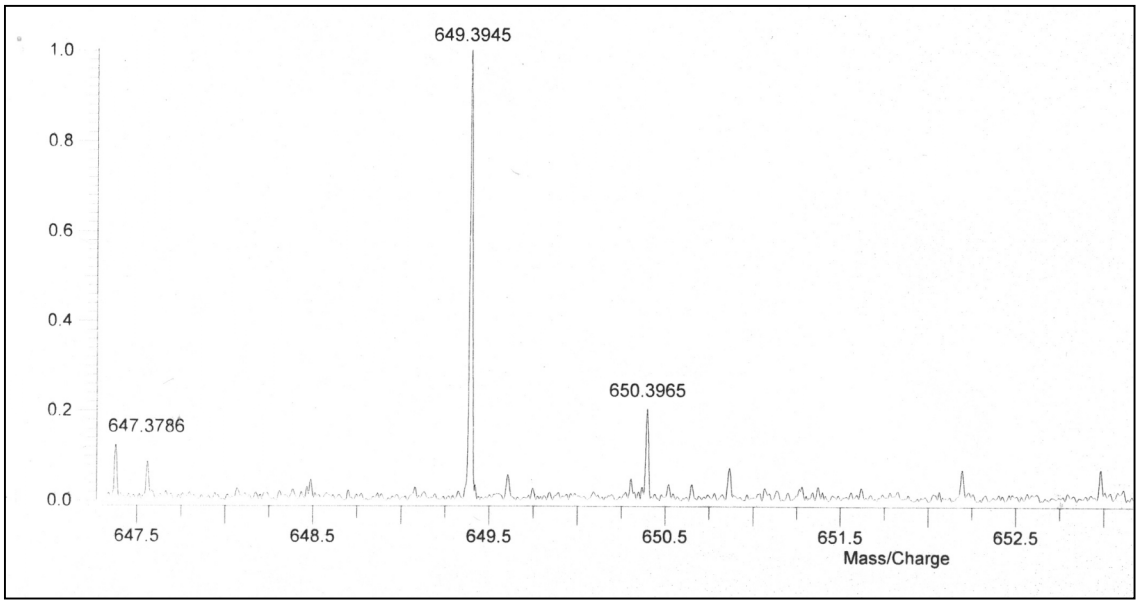


Figure A.47. HR-ESI-MS negative mode spectra for *S. radula* compound **9.**

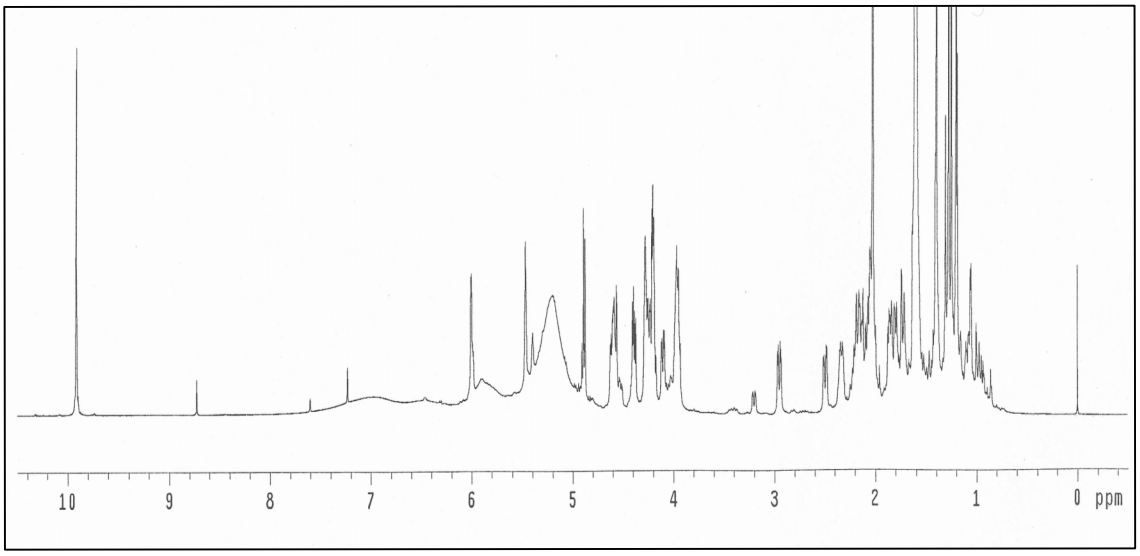


Figure A.48. ^{1}H-NMR spectra for *S. radula* compound **9.**

.

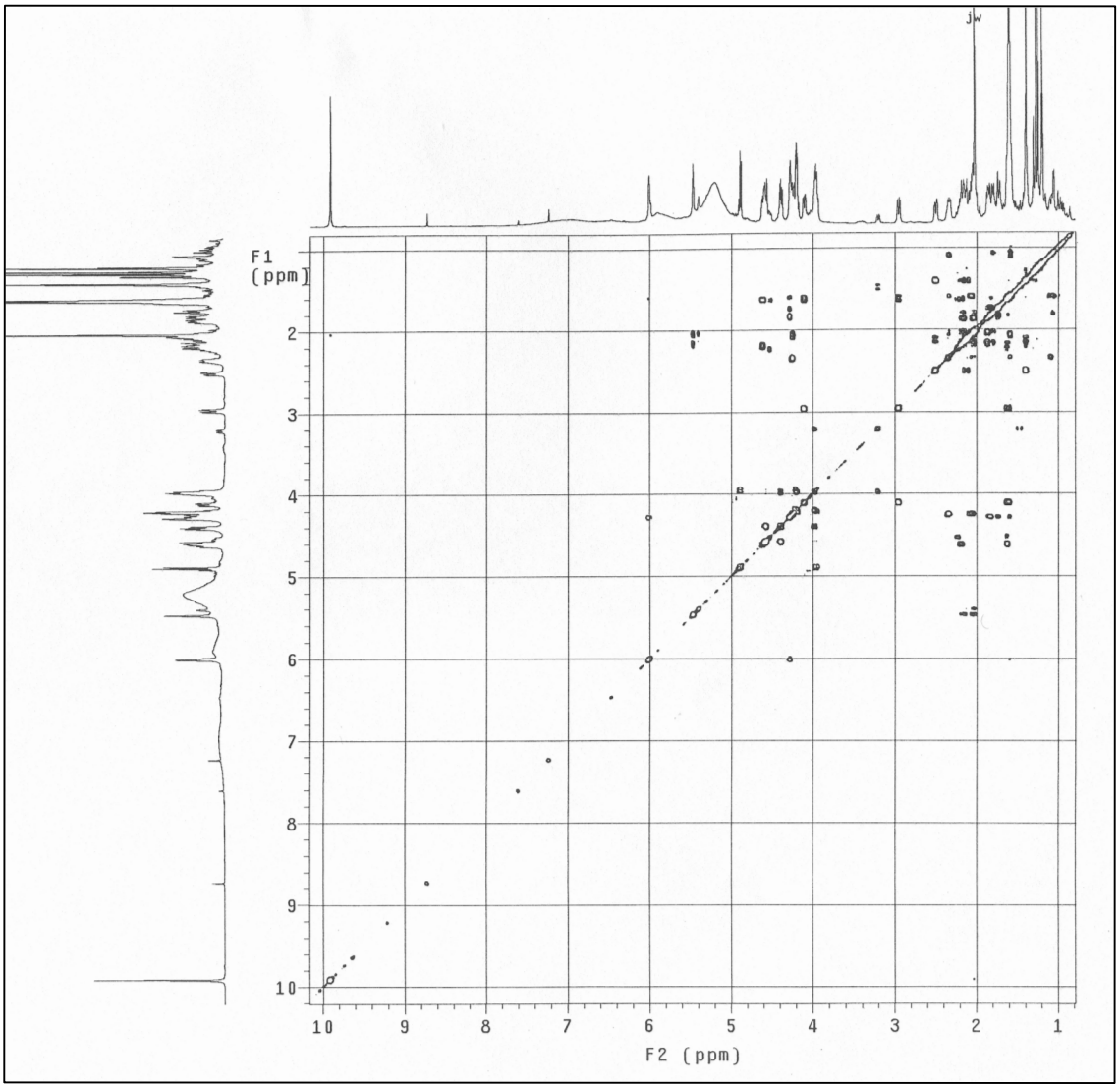


Figure A.49. COSY spectra for *S. radula* compound **9.**

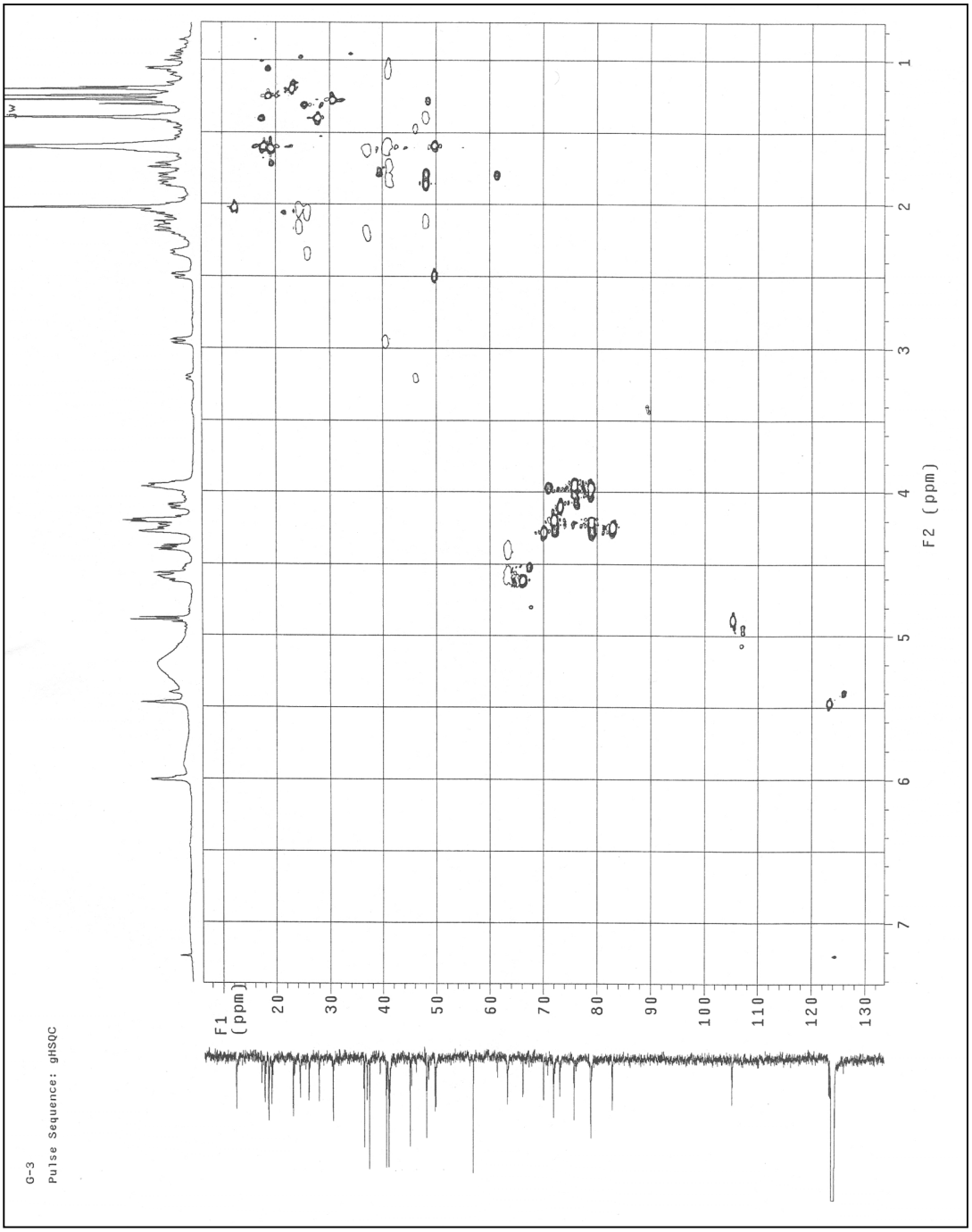


Figure A.50. HSQC spectra for *S. radula* compound **9**.

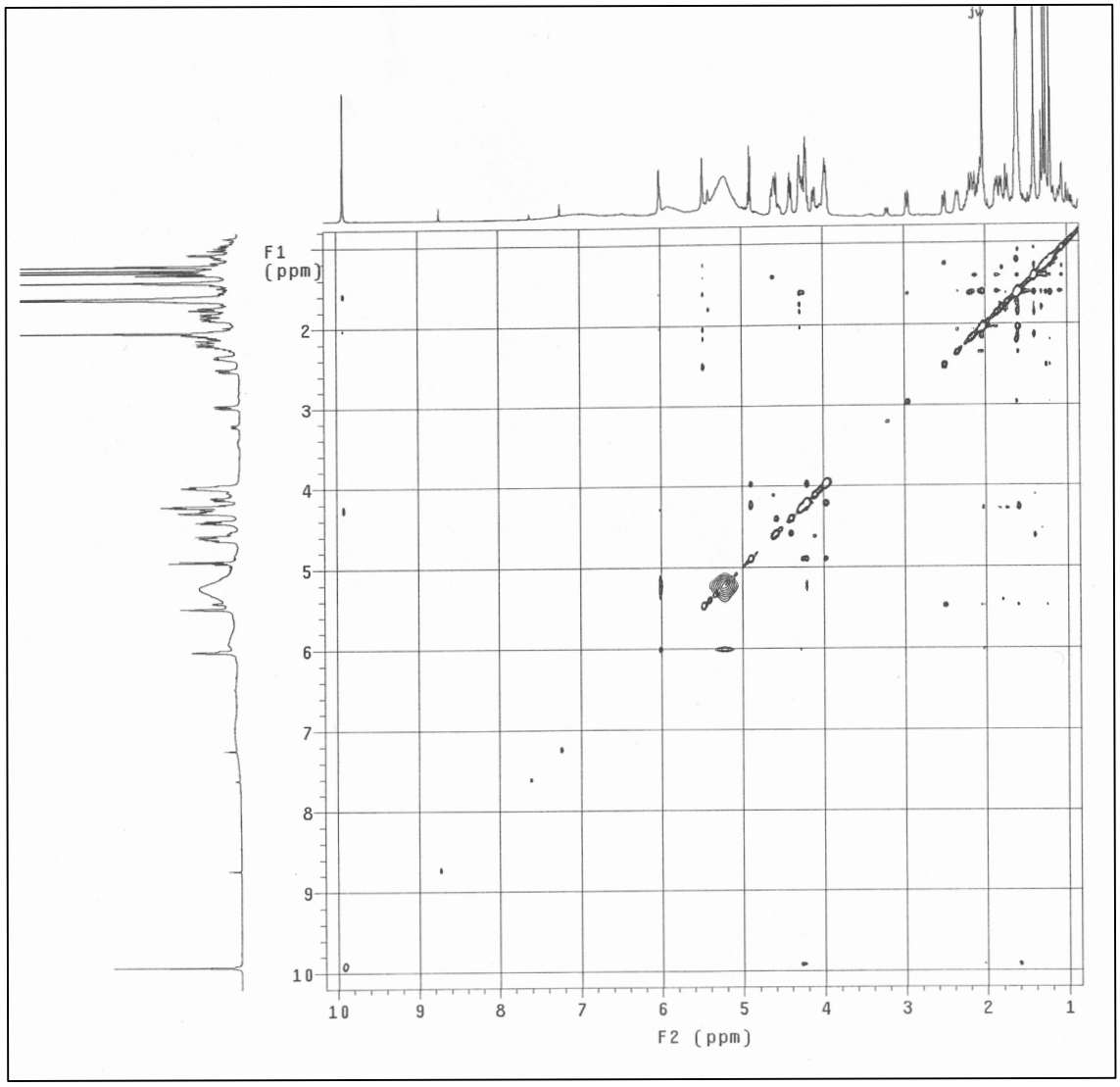


Figure A.51. NOESY spectra for *S. radula* compound **9**.

Position	1		2		3		4		5	
	δ_C	δ_H (J in Hz)	δ_C	δ_H (J in Hz)	δ_C	δ_H (J in Hz)	δ_C	δ_H (J in Hz)	δ_C	δ_H (J in Hz)
1	40.5	1.54, 1.03, *m*	41.4	1.53, 0.92, *m*	42.0	1.63, 1.07, *m*	41.4	1.57, 0.97, *m*	41.2	1.57, 1.04, *m*
2	25.6	2.30, 2.03, *m*	27.0	2.30, 2.06, *m*	26.7	1.93, 1.78, *m*	27.0	2.37, 2.10, *m*	26.9	2.36, 2.10, *m*
3	82.3	4.22, *dd* (11.0, 4.5)	89.4	3.38, *dd* (11.0, 4.5)	87.1	4.00, *dd* (11.0, 4.5)	89.3	3.47, *dd* (11.0, 4.5)	89.2	3.45, *dd* (11.0, 4.5)
4	56.4		40.7		55.3		40.7		40.7	
5	49.3	1.50, *br s*	56.3	0.85, *br s*	53.6	1.52, *br s*	56.3	0.94, br s	56.3	0.95, *br s*
6	69.4	4.13, *br s*	67.1	4.79, *br s*	71.4	4.10, *br s*	67.3	4.81, br s	67.3	4.81, *br s*
7	40.7	1.73, 1.66, *m*	41.4	1.83, 1.53, *m*	41.3	1.76, 1.53, *m*	41.4	1.90, 1.87, *m*	41.2	1.88, 1.87, *m*
8	40.2		39.9		40.9		39.9		40.6	
9	47.6	1.80, *m*	47.9	1.66, *m*	48.9	1.65, *m*	47.9	1.72, *m*	47.8	1.80, *m*
10	35.9		36.5		37.0		36.6		36.7	
11	23.8	2.11, *dd* (12.0, 3.5) 1.97, *dd* (12.0, 3.5)	23.8	2.17, *dd* (12.0, 3.5) 2.02, *dd* (12.0, 3.5)	24.7	2.06, *dd* (12.0, 3.5) 1.94, *dd* (12.0, 3.5)	23.9	2.18, *dd* (12.0, 3.5) 2.03, *dd* (12.0, 3.5)	24.0	2.20, *dd* (12.0, 3.5) 2.06, *dd* (12.0, 3.5)
12	122.6	5.38, *t* (3.5)	125.5	5.38, *t* (3.5)	123.4	5.28, t (3.5)	126.0	5.42, *t* (3.5)	123.0	5.49, *t* (3.5)
13	143.5		138.5		144.4		138.6		143.1	
14	44.5		44.7		45.3		44.8		44.5	
15	34.8	2.20, *dd* (12.0, 11.5) 2.03, *dd* (12.0, 4.5)	36.5	2.24, *dd* (13.0, 11.5) 1.66, *dd* (13.0, 4.5)	35.4	1.83, *dd* (12.0, 11.5) 1.57, *dd* (12.0, 4.5)	37.1	2.29, 1.69, *m*	36.6	2.25, *dd* (12.0, 11.5) 1.69, *dd* (12.0, 4.5)
16	76.4	4.58, *dd* (11.5, 4.5)	65.7	4.55, *dd* (11.5, 4.5)	77.6	4.15, *dd* (11.5, 4.5)	67.0	4.56, *dd* (11.5, 4.5)	65.8	4.65, *dd* (11.5, 4.5)
17	38.3		39.3		38.9		40.8		39.7	
18	49.9	2.31, *m*	61.2	1.67, *d* (11.5)	51.0	2.15, *m*	61.0	1.83, *d* (11.5)	49.3	2.51. m
19	47.2	1.97, 1.20, *m*	40.0	1.56, *m*	48.0	1.77, 1.05, *m*	39.0	1.82, m	47.8	2.16, 1.42, *m*
20	31.2		39.9	0.98, *m*	31.9		48.2	1.29, *m*	36.9	
21	35.2	1.80, 1.27, *m*	31.3	1.54, 1.46, *m*	35.8	1.43, 1.14, *m*	70.5	3.97, *dd* (11.0, 4.0)	72.7	4.12, *dd* (11.0, 4.0)
22	30.7	2.80, 1.33, *m*	36.0	2.60, 1.17, *m*	31.2	2.17, 1.13, *m*	45.9	3.21, *dd* (11.0, 4.0) 1.50, *dd* (11.0, 11.0)	40.2	2.98, *dd* (11.0, 4.0) 1.63, *dd* (11.0, 11.0)
23	207.6	9.87, *s*	28.1	1.54, *s*	183.1	–	28.2	1.54, *s*	28.1	1.54, *s*
24	11.9	2.00, *s*	18.6	1.84, *s*	13.9	1.52, *s*	18.7	1.73, *s*	18.7	1.74, *s*
25	18.4	1.56, *s*	17.4	1.62, *s*	17.7	1.36, *s*	17.4	1.62, *s*	17.2	1.62, *s*
26	17.3	1.51, *s*	18.8	1.66, *s*	18.8	1.27, *s*	18.8	1.65, *s*	18.6	1.64, *s*
27	27.4	1.44, *s*	25.0	1.33, *s*	27.7	1.21, *s*	25.0	1.35, *s*	27.4	1.44, *s*
28	23.2	1.15, *s*	23.0	1.15, *s*	23.3	0.87, *s*	22.9	1.22, *s*	22.6	1.23, *s*
29	33.4	0.95, *s*	18.0	0.99, *d* (6.5)	33.8	0.91, *s*	18.2	1.08, *d* (6.5)	30.1	1.29, *s*
30	24.2	0.98, *s*	21.7	0.98, *d* (6.5)	24.4	0.93, *s*	16.8	1.41, *d* (6.5)	18.0	1.26, *s*
	Glc at C-3		Glc at C-3		Glc at C-3		Glc at C-3		Glc at C-3	
1′	104.9	4.89, *d* (7.5)	105.0	4.98, *d* (7.5)	105.3	4.30, *d* (7.5)	107.0	5.01, *d* (7.5)	106.9	5.00, *d* (7.5)
2′	75.4	3.95, *dd* (9.0, 7.5)	84.6	4.24, *dd* (9.0, 7.5)	75.5	3.14, *dd* (9.0, 7.5)	75.9	4.10, *dd* (9.0, 7.5)	75.9	4.10, *dd* (9.0, 7.5)
3′	78.6	4.19, *dd* (9.0, 9.0)	78.4	4.32, *dd* (9.0, 9.0)	77.6	3.30, *dd* (9.0, 9.0)	78.8	4.29, *dd* (9.0, 9.0)	78.8	4.28, *dd* (9.0, 9.0)
4′	71.6	4.18, *dd* (9.0, 9.0)	71.6	4.19, *dd* (9.0, 9.0)	71.6	3.30, *dd* (9.0, 9.0)	72.0	4.26, *dd* (9.0, 9.0)	71.9	4.27, *dd* (9.0, 9.0)
5′	78.5	3.96, *m*	78.0	3.95, *m*	78.0	3.27, *m*	78.4	4.06, *m*	78.3	4.05, m
6′	62.9	4.40, *dd* (12.0, 5.0) 4.58, *dd* (12.0, 2.5)	62.9	4.38, *dd* (12.0, 5.0) 4.58, *dd* (12.0, 2.5)	62.8	3.67, *dd* (12.0, 5.0) 3.82, *dd* (12.0, 2.5)	63.2	4.47, *dd* (12.0, 5.0) 4.64, *dd* (12.0, 2.5)	63.1	4.45, *dd* (12.0, 5.0) 4.3, *dd* (12.0, 2.5)

(*continued on next page*)

Table A.1 ^{13}C and ^{1}H-NMR spectroscopic data for compounds **1-5** (**1,2,4,5** in $C_5D_5N_5$; **3** in CD_3OD).

Position	**1**		**2**		**3**		**4**		**5**	
	δ_C	δ_H (*J* in Hz)	δ_C	δ_H (*J* in Hz)	δ_C	δ_H (*J* in Hz)	δ_C	δ_H (*J* in Hz)	δ_C	δ_H (*J* in Hz)
	Glc at C-16		Gal at C-2′glc		Glc at C-16					
1″	106.3	5.06, *d* (7.5)	107.3	5.25, *d* (7.5)	106.1	4.30, *d* (7.5)				
2″	75.8	4.03, *dd* (9.0, 7.5)	74.8	4.61, *dd* (8.6, 7.5)	75.7	3.16, *dd* (9.0, 7.5)				
3″	78.8	4.26, *dd* (9.0, 9.0)	75.0	4.18, *dd* (8.6, 4.0)	77.9	3.30, *dd* (9.0, 9.0)				
4″	71.9	4.25, *dd* (9.0, 9.0)	69.6	4.70, *dd* (4.0, 2.0)	71.8	3.30, *dd* (9.0, 9.0)				
5″	78.1	3.99, *m*	76.9	4.07, *m*	78.3	3.34, *m*				
6″	63.0	4.36, *dd* (12.0, 5.0) 4.48, *dd* (12.0, 2.5)	61.4	4.41, *dd* (12.0, 4.5) 4.58, *dd* (12.0, 2.0)	62.9	3.68, *dd* (12.0, 5.0) 3.82, *dd* (12.0, 2.5)				

Table A.1 ***continued*** ^{13}C and ^{1}H-NMR spectroscopic data for compounds **1-5** (**1,2,4,5** in $C_5D_5N_5$; **3** in CD_3OD).

Position	6		7		8		9	
	δ_C	δ_H (J in Hz)	δ_C	δ_H (J in Hz)	δ_C	δ_H (J in Hz)	δ_C	δ_H (J in Hz)
1	41.8	1.57, 0.96, *m*	41.5	1.75, 1.57, *m*	40.7	1.80, 1.97, m	40.6	1.58, 1.05, *m*
2	26.4	1.94, 1.83, *m*	27.2	1.84, 1.78, *m*	25.7	2.35, 2.16, *m*	25.5	2.33, 2.05, *m*
3	83.3	3.59, *dd* (11.0, 4.5)	90.5	3.13, *dd* (11.0, 4.5)	81.8	4.29, *dd* (11.0, 4.5)	82.5	4.24, *dd* (11.0, 4.5)
4	44.8		40.3	–	44.2		56.3	
5	48.2	1.23, *br s*	57.3	0.76, *br s*	47.9	1.83, *br s*	49.3	1.59, *br s*
6	68.4	4.45, *br s*	68.5	4.51, *br s*	66.7	4.99, *br s*	69.6	4.27, *br s*
7	41.0	1.83, 1.54, *m*	42.1	1.57, 0.97, *m*	40.4	1.59, 1.09, *m*	40.4	1.83, 1.70, *m*
8	41.2		41.0	–	39.4		40.5	
9	47.8	1.63, *m*	49.0	1.58, *m*	47.4	1.74, *m*	47.6	1.84, *m*
10	37.0		37.2	–	36.1		35.8	
11	24.6	2.04, *dd* (12.0, 3.5) 1.93, *dd* (12.0, 3.5)	24.6	2.05, *dd* (12.0, 3.5) 1.89, *dd* (12.0, 3.5)	23.6	2.17, *dd* (12.0, 3.5) 2.01, *dd* (12.0, 3.5)	23.9	2.15, *dd* (12.0, 3.5) 2.03, *dd* (12.0, 3.5)
12	124.2	5.29, *t* (3.5)	123.4	5.25, *t* (3.5)	122.3	5.38, *t* (3.5)	123.8	5.47 t (3.5)
13	143.2		144.3	–	143.2		143.1	
14	45.3		45.2	–	44.2		44.5	
15	35.2	1.83, *dd* (12.0, 11.5) 1.56 *dd* (12.0, 4.5)	35.5	1.84, *dd* (12.0, 11.5) 1.57, *dd* (12.0, 4.5)	34.7	2.23, *dd* (12.0, 11.5) 1.96, *dd* (12.0, 4.5)	36.3	2.20, *dd* (12.0, 11.5) 1.62, *dd* (12.0, 4.5)
16	78.0	4.10, *dd* (12.0, 4.0)	77.2	4.16, *dd* (11.5, 4.5)	75.9	4.54, *dd* (11.5, 4.5)	65.7	4.61, *dd* (11.5, 4.5)
17	40.3		38.9	–	37.9		40.0	
18	50.2	2.18, *m*	51.0	2.14, *m*	49.6	2.32, *m*	49.2	2.50, *m*
19	47.4	1.84, 1.14, *m*	48.2	1.76, 1.04, *m*	46.9	1.59, 1.16, *m*	47.7	2.14, 1.39, *m*
20	37.0		31.8		30.8		36.8	
21	74.2	3.55, *dd* (11.8, 4.5)	35.9	1.38, 1.13, *m*	34.9	1.66, 1.27, m	72.6	4.10, *dd* (11.8, 4.5)
22	40.2	1.82, *dd* (11.0, 4.0) 1.53, *dd* (11.0, 4.0)	30.8	2.28, 1.06, *m*	29.9	2.84, *dd* (11.0, 4.0) 1.34, *dd* (11.0, 4.0)	40.7	2.95, *dd* (11.0, 4.0) 1.64, *dd* (11.0, 4.0)
23	64.6	3.48, *dd* (12.0) 3.74, *dd* (12.0)	28.2	1.13, *s*	64.0	4.00, *dd* (12.0) 4.50, *dd* (12.0)	207.7	9.91, *s*
24	15.0	1.09, *s*	18.5	1.23, *s*	15.0	1.62, *s*	11.9	2.01, *s*
25	18.0	1.27, *s*	17.5	1.33, *s*	17.4	1.64, *s*	17.2	1.59, *s*
26	18.8	1.35, *s*	18.8	1.27, *s*	18.0	1.54, *s*	18.5	1.62, *s*
27	27.5	1.19, *s*	27.5	1.18, *s*	27.0	1.38, *s*	27.4	1.39, *s*
28	23.3	0.89, *s*	23.2	0.86, *s*	22.8	1.22, *s*	22.5	1.20, *s*
29	29.8	0.97, *s*	31.8	0.90, *s*	33.0	0.86, *s*	30.0	1.27, *s*
30	17.9	0.88, *s*	24.4	0.91, *s*	23.9	0.94, *s*	17.9	1.26, *s*
	Glc at C-3		Glc at C-3		Glc at C-3		Glc at C-3	
1′	105.7	4.40, *d* (7.5)	106.6	4.32, *d* (7.5)	105.2	5.17, *d* (7.5)	104.8	4.90, *d* (7.5)
2′	75.6	3.20, *dd* (9.0, 7.5)	75.6	3.20, *dd* (9.0, 7.5)	75.4	4.05, *dd* (9.0, 7.5)	75.3	3.95, *dd* (9.0, 7.5)
3′	78.3	3.35, *dd* (9.0, 9.0)	78.2	3.33, *dd* (9.0, 9.0)	78.2	4.19, *dd* (9.0, 9.0)	78.5	4.21, *dd* (9.0, 9.0)
4′	71.7	3.33, *dd* (9.0, 9.0)	71.6	3.31, *dd* (9.0, 9.0)	71.1	4.20, *dd* (9.0, 9.0)	71.5	4.27, *dd* (9.0, 9.0)
5′	77.5	3.27, *m*	77.6	3.26, *m*	77.8	3.90, *m*	78.5	3.96, *m*
6′	62.8	3.85, *dd* (12.0, 2.5) 3.67, *dd* (12.0, 5.0)	62.8	3.86, *dd* (12.0, 2.5) 3.65, *dd* (12.0, 5.0)	62.3	4.50, *dd* (12.0, 2.5) 4.33, *dd* (12.0, 5.0)	62.8	4.56, *dd* (12.0, 2.5) 4.40, *dd* (12.0, 5.0)

(*continued on next page*)

Table A.2 . ^{13}C and ^{1}H-NMR spectroscopic data for compounds **6-9** (**6,7** in CD_3OD; **8,9** in $C_5D_5N_5$).

Position	**6**		**7**		**8**		**9**	
	δ_C	δ_H (*J* in Hz)	δ_C	δ_H (*J* in Hz)	δ_C	δ_H (*J* in Hz)	δ_C	δ_H (*J* in Hz)
	Glc at C-16		Glc at C-16		Glc at C-16			
1″	105.5	4.37, *d* (7.5)	104.5	4.43, *d* (7.5)	103.9	4.99, *d* (7.5)		
2″	75.6	3.20, *dd* (9.0, 7.5)	83.5	3.40, *dd* (9.0, 7.5)	82.9	4.09, *dd* (9.0, 7.5)		
3″	78.3	3.35, *dd* (9.0, 9.0)	78.0	3.55, *dd* (9.0, 9.0)	77.8	4.28, *dd* (9.0, 9.0)		
4″	71.5	3.27, *dd* (9.0, 9.0)	71.6	3.29, *dd* (9.0, 9.0)	71.2	4.20, *dd* (9.0, 9.0)		
5″	77.6	3.33, *m*	77.3	3.26, *m*	77.2	3.87, *m*		
6″	62.7	3.85, *dd* (12.0, 2.5) 3.67, *dd* (12.0, 5.0)	62.7	3.86, *dd* (12.0, 2.5) 3.65, *dd* (12.0, 5.0)	62.3	4.50, *dd* (12.0, 2.5) 4.33, *dd* (12.0, 5.0)		
			Ara at C-2glcI′		Ara at C-2glcI			
1‴			106.4	4.48, *d* (6.5)	106.0	5.18, *d* (6.5)		
2‴			73.6	3.65, *dd* (8.5, 6.5)	73.2	4.55, *dd* (8.5, 6.5)		
3‴			74.0	3.55, *dd* (8.5, 3.0)	73.8	4.19, *dd* (8.5, 3.0)		
4‴			69.5	3.80, *m*	68.7	4.32, *m*		
5‴			67.3	3.84, *dd* (12.0, 2.0) 3.50, *dd* (12.0, 3.0)	66.6	4.02, *dd* (12.0, 2.0) 3.74, *dd* (12.0, 3.0)		

Table A.2 *continued* ^{13}C and ^{1}H-NMR spectroscopic data for compounds **6-9** (**6,7** in CD_3OD; **8,9** in $C_5D_5N_5$).

Appendix B

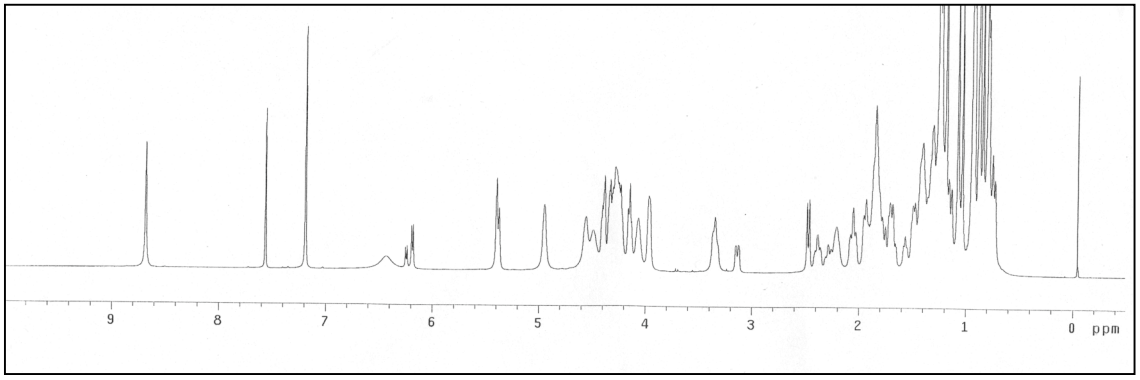


Figure B.1. ^{1}H-NMR spectra for *S. integrifolium* compound **1.**

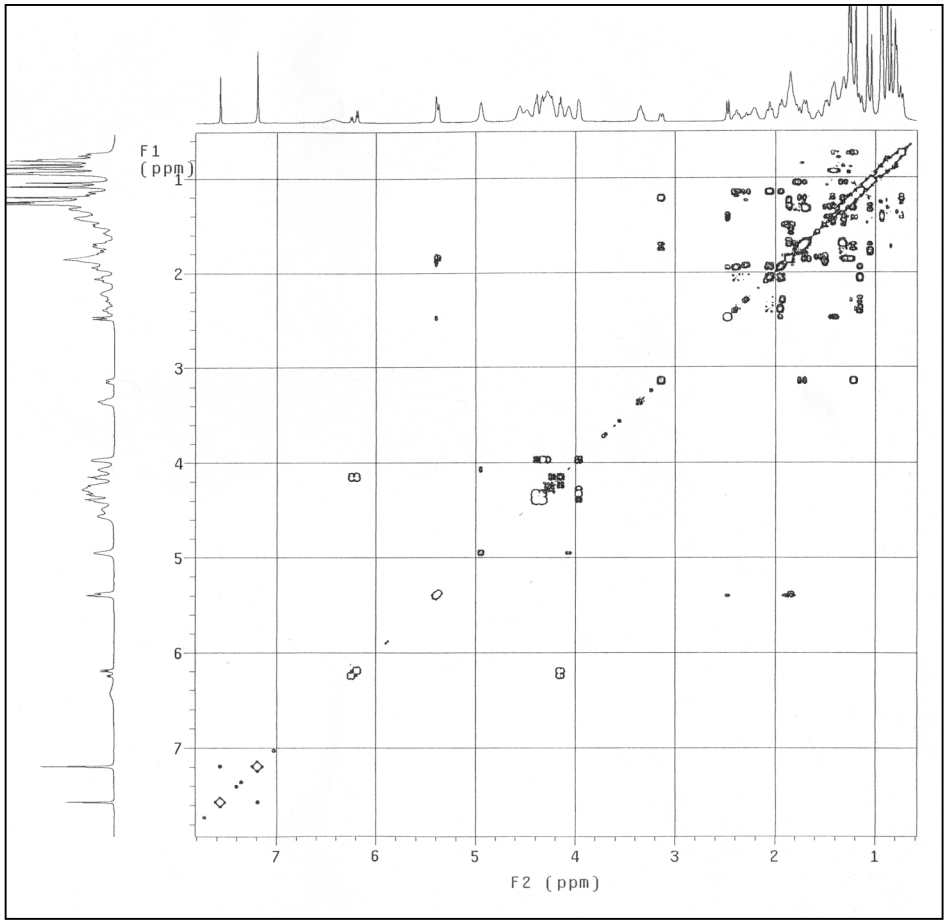


Figure B.2. COSY spectra for *S. integrifolium* compound **1.**

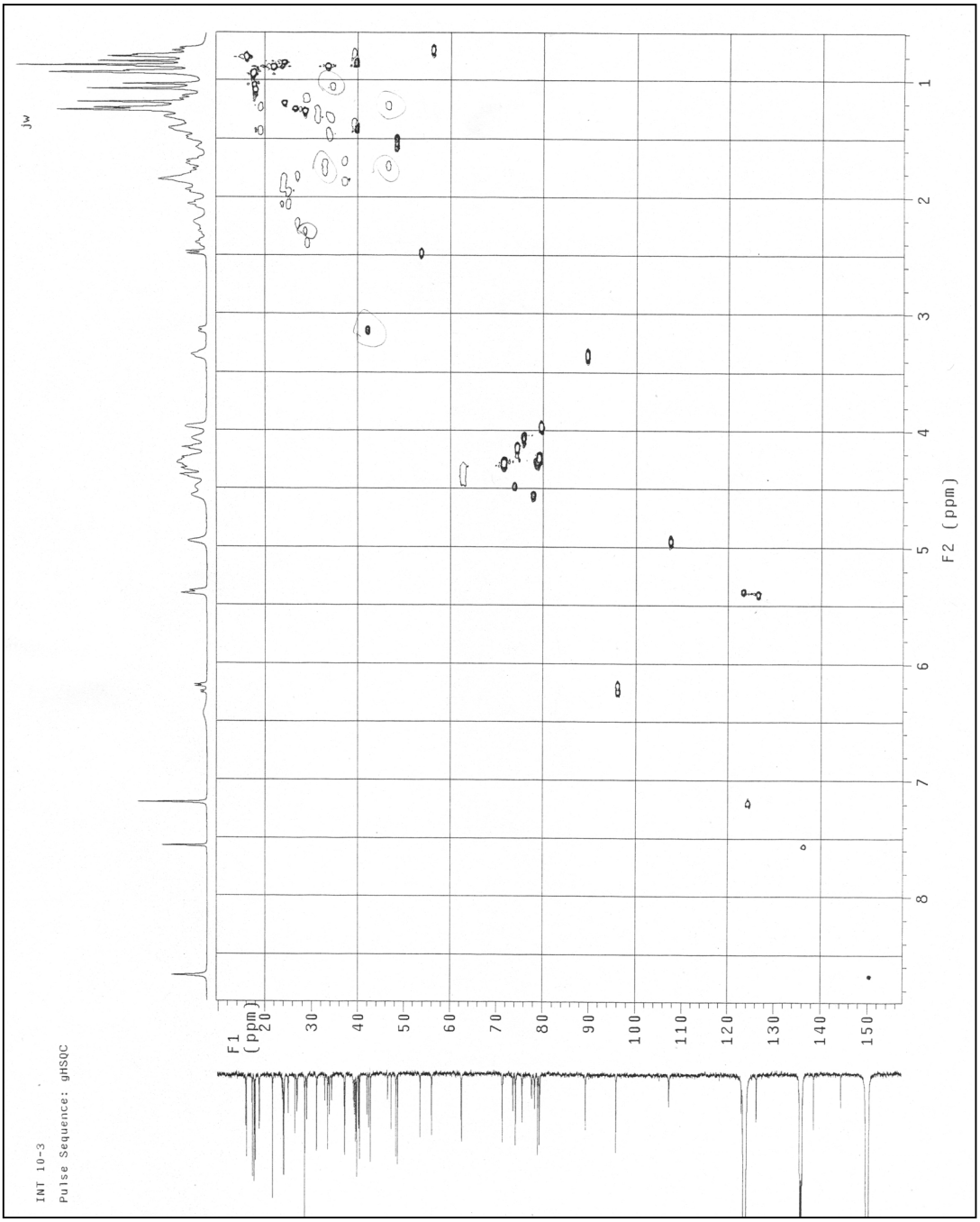


Figure B.3. HSQC spectra for *S. integrifolium* compound **1.**

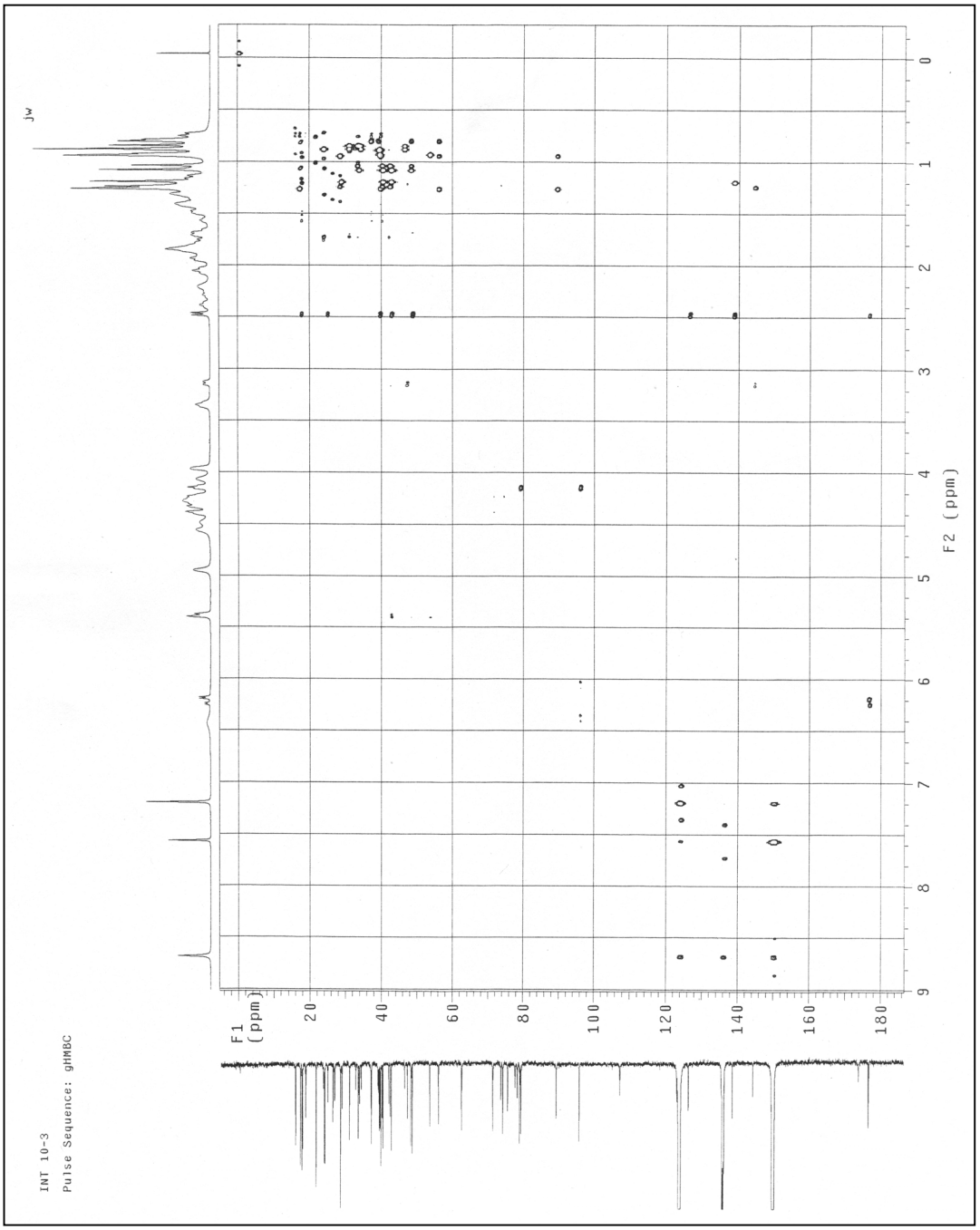


Figure B.4. HMBC spectra for *S. integrifolium* compound **1.**

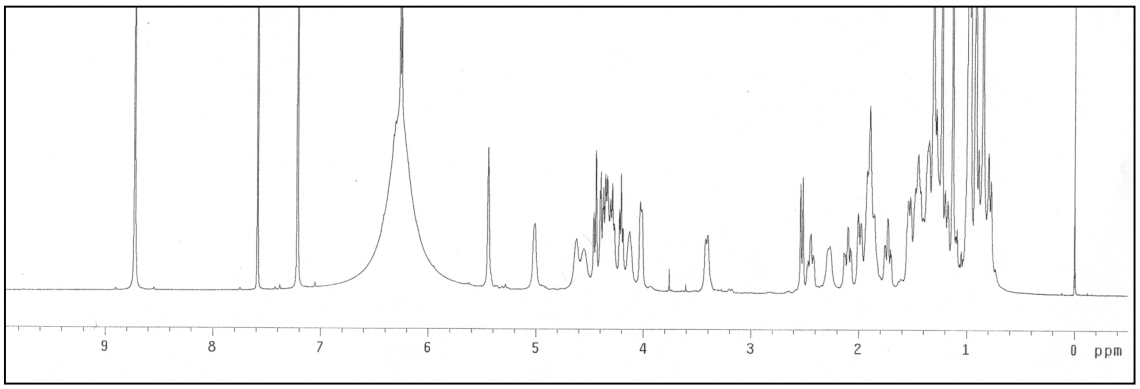


Figure B.5. ^{1}H-NMR spectra for *S. integrifolium* compound **2.**

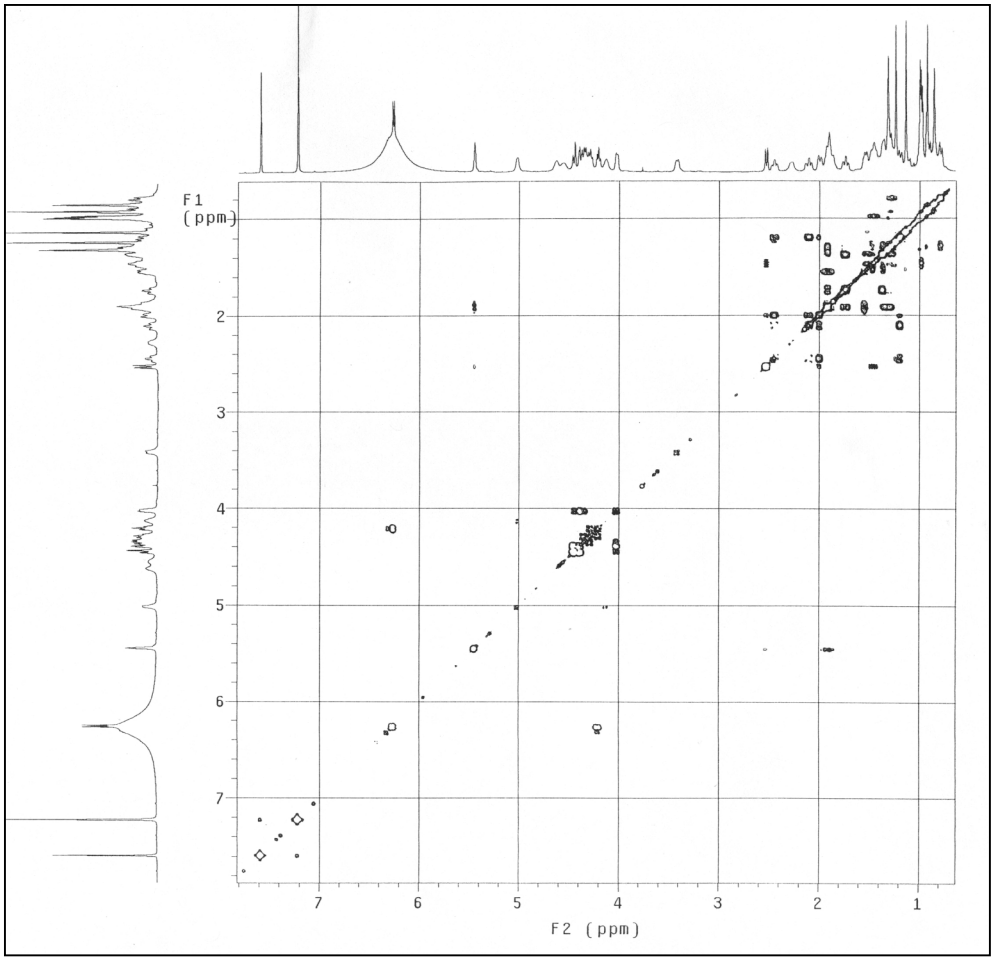


Figure B.6. COSY spectra for *S. integrifolium* compound **2.**

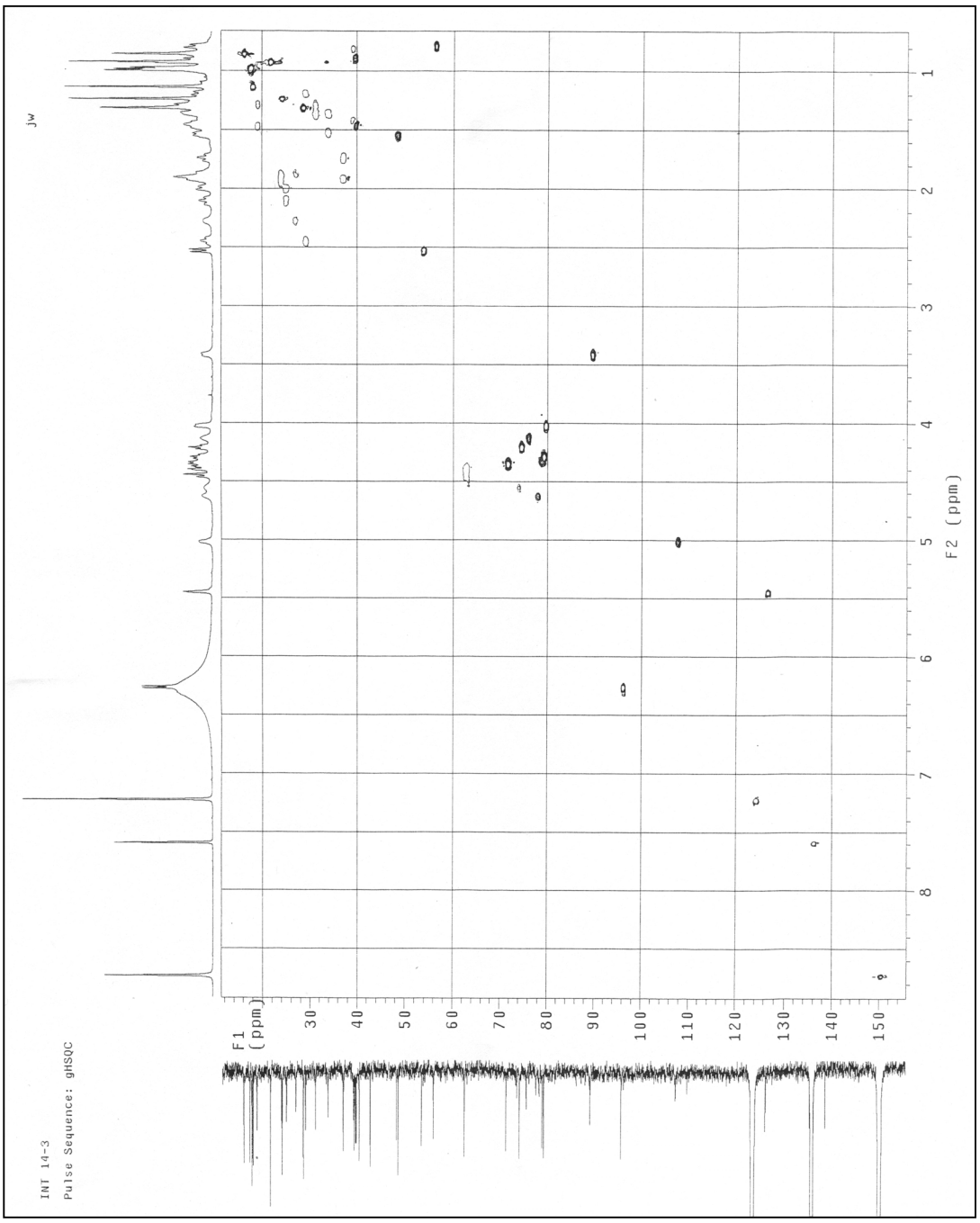


Figure B.7. HSQC spectra for *S. integrifolium* compound **2.**

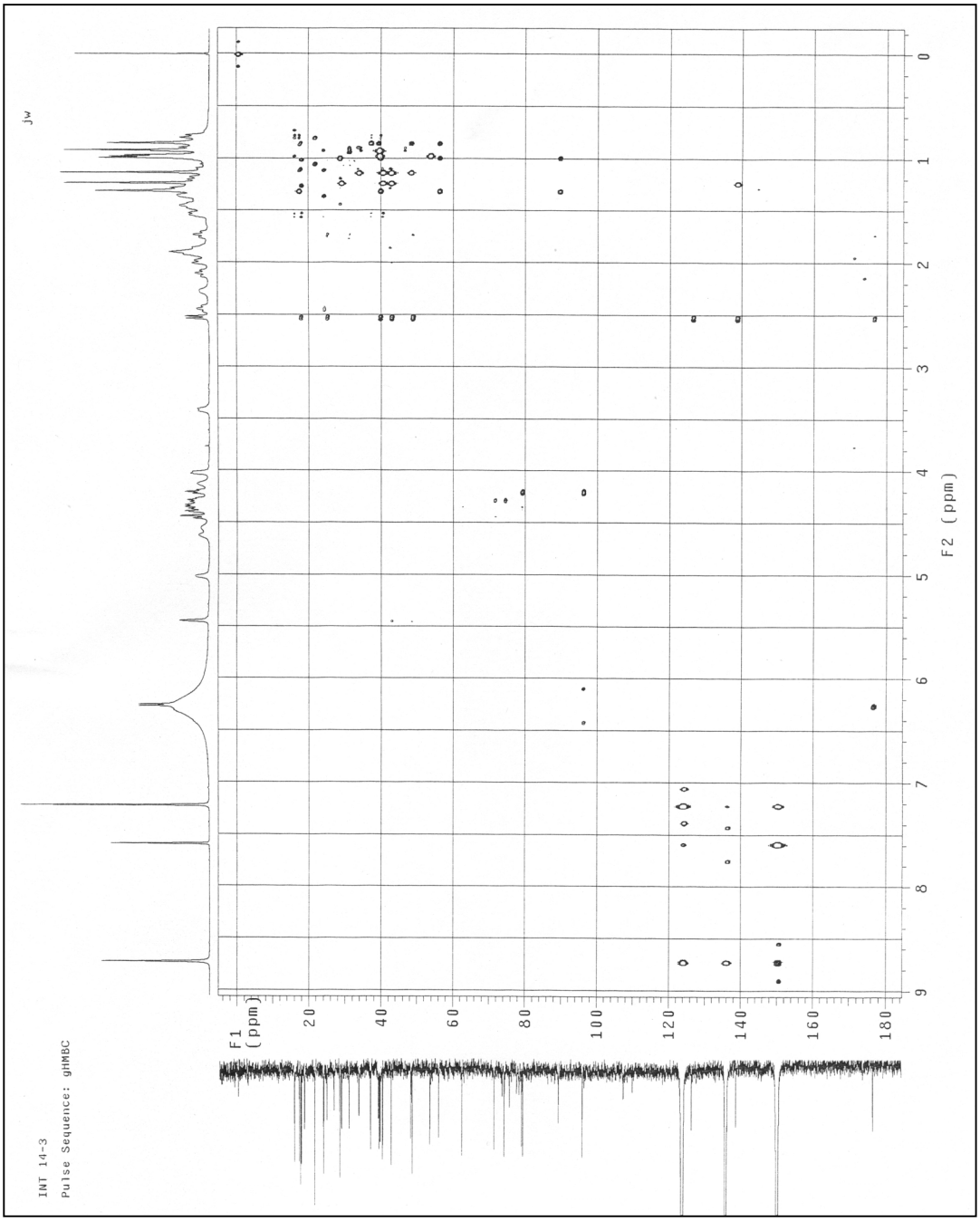


Figure B.8. HMBC spectra for *S. integrifolium* compound **2.**

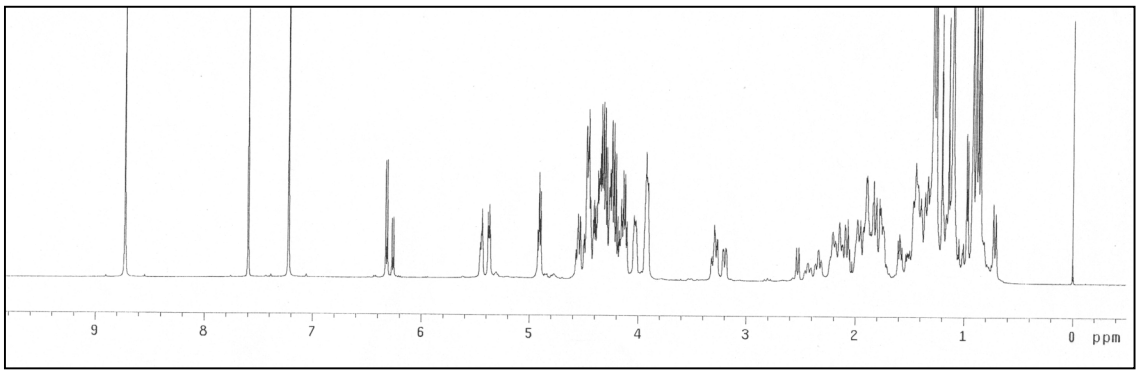


Figure B.9. [1]H-NMR spectra for *S. integrifolium* compound **3.**

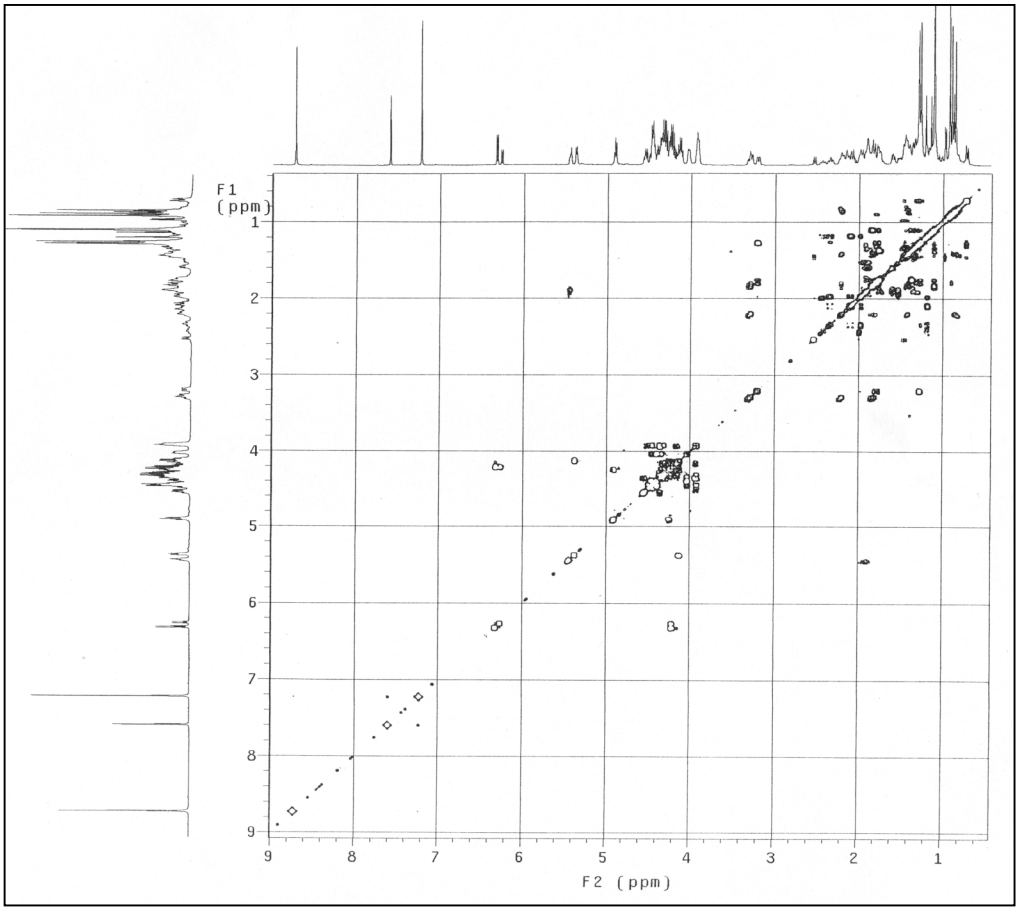


Figure B.10. COSY spectra for *S. integrifolium* compound **3.**

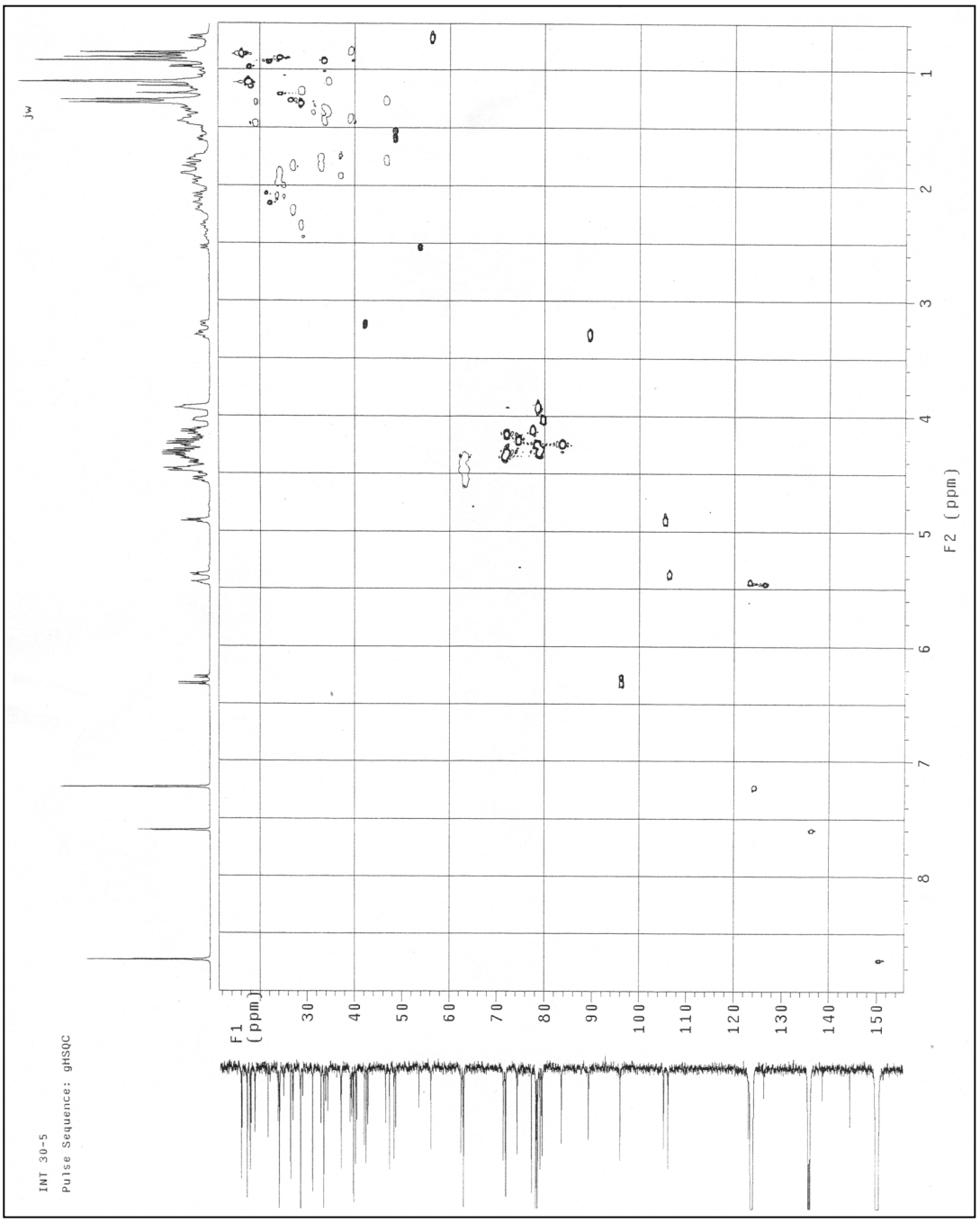


Figure B.11. HSQC spectra for *S. integrifolium* compound **3.**

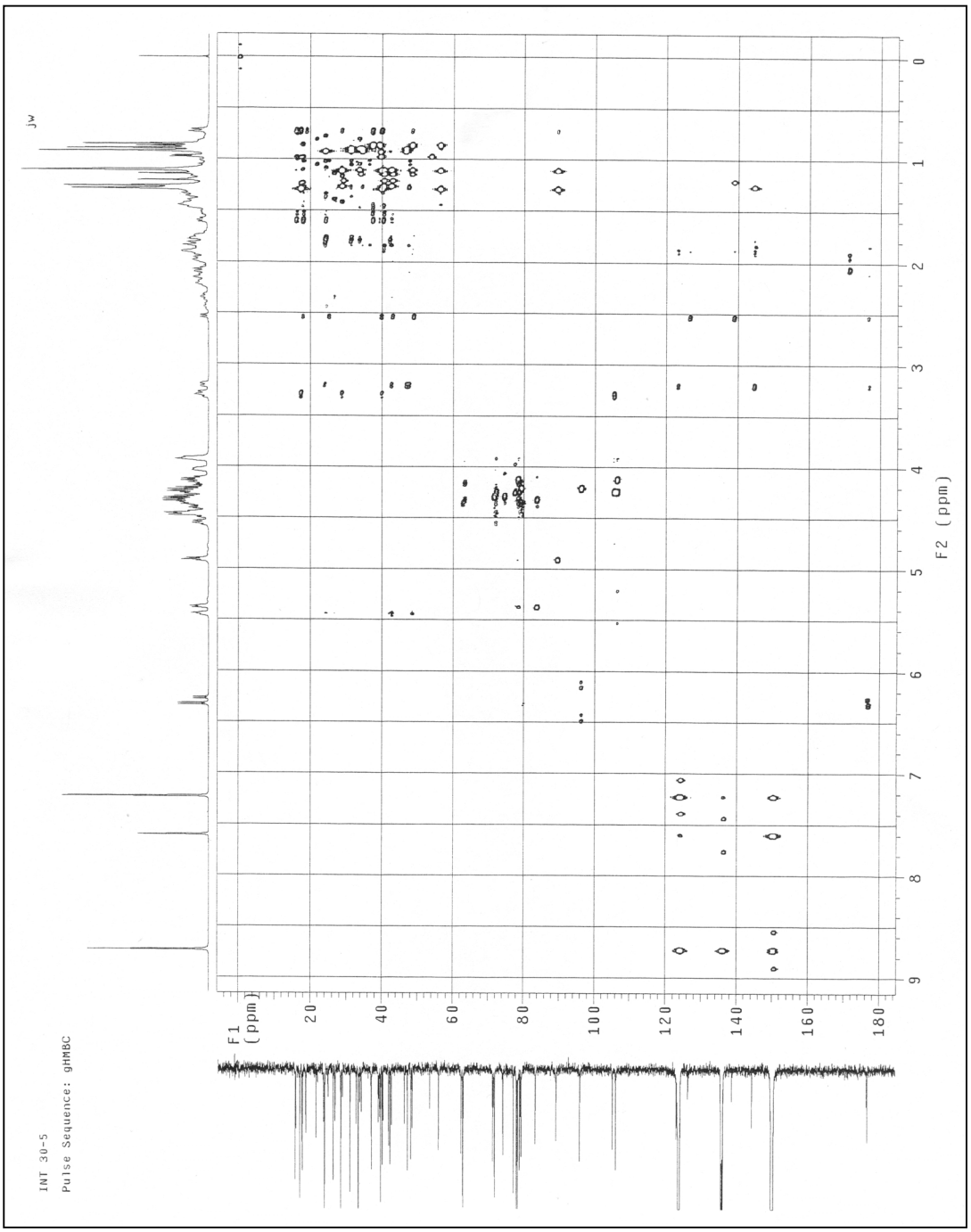


Figure B.12. HMBC spectra for *S. integrifolium* compound **3.**

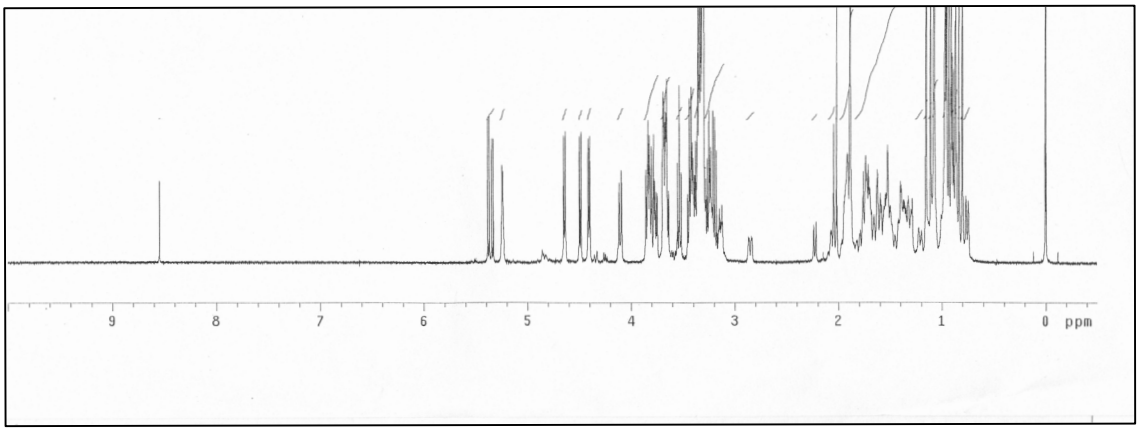


Figure B.13. ^{1}H-NMR spectra for *S. integrifolium* compound **4.**

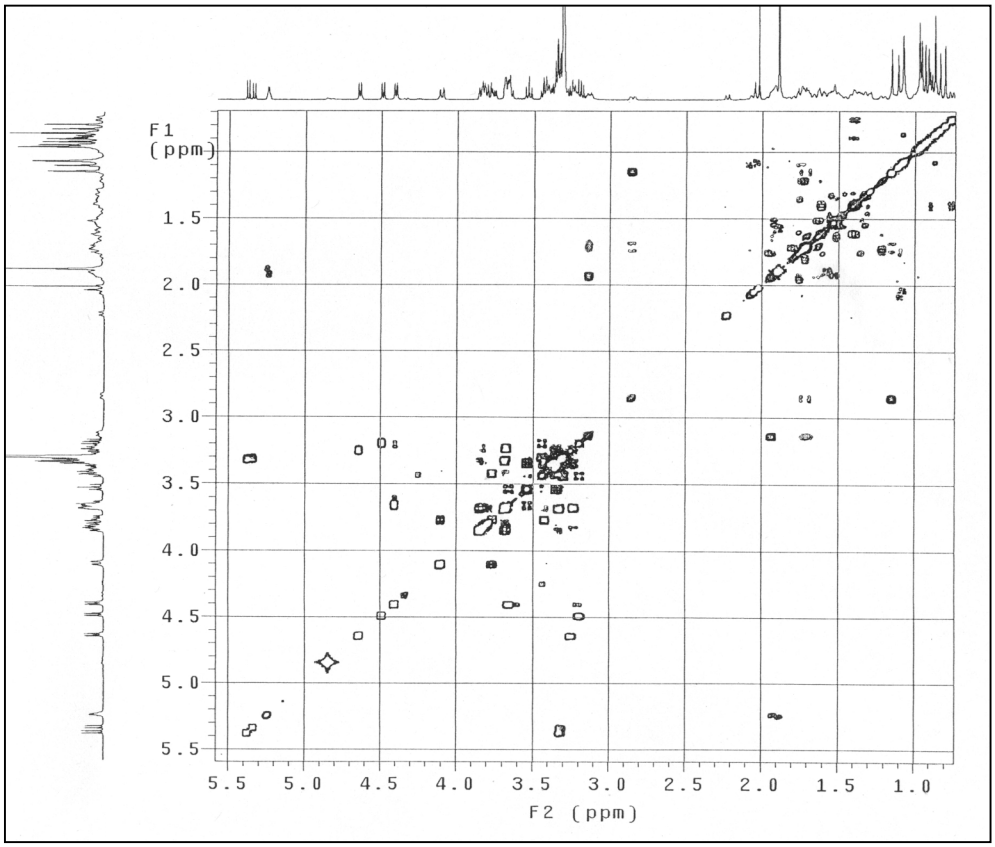


Figure B.14. COSY spectra for *S. integrifolium* compound **4.**

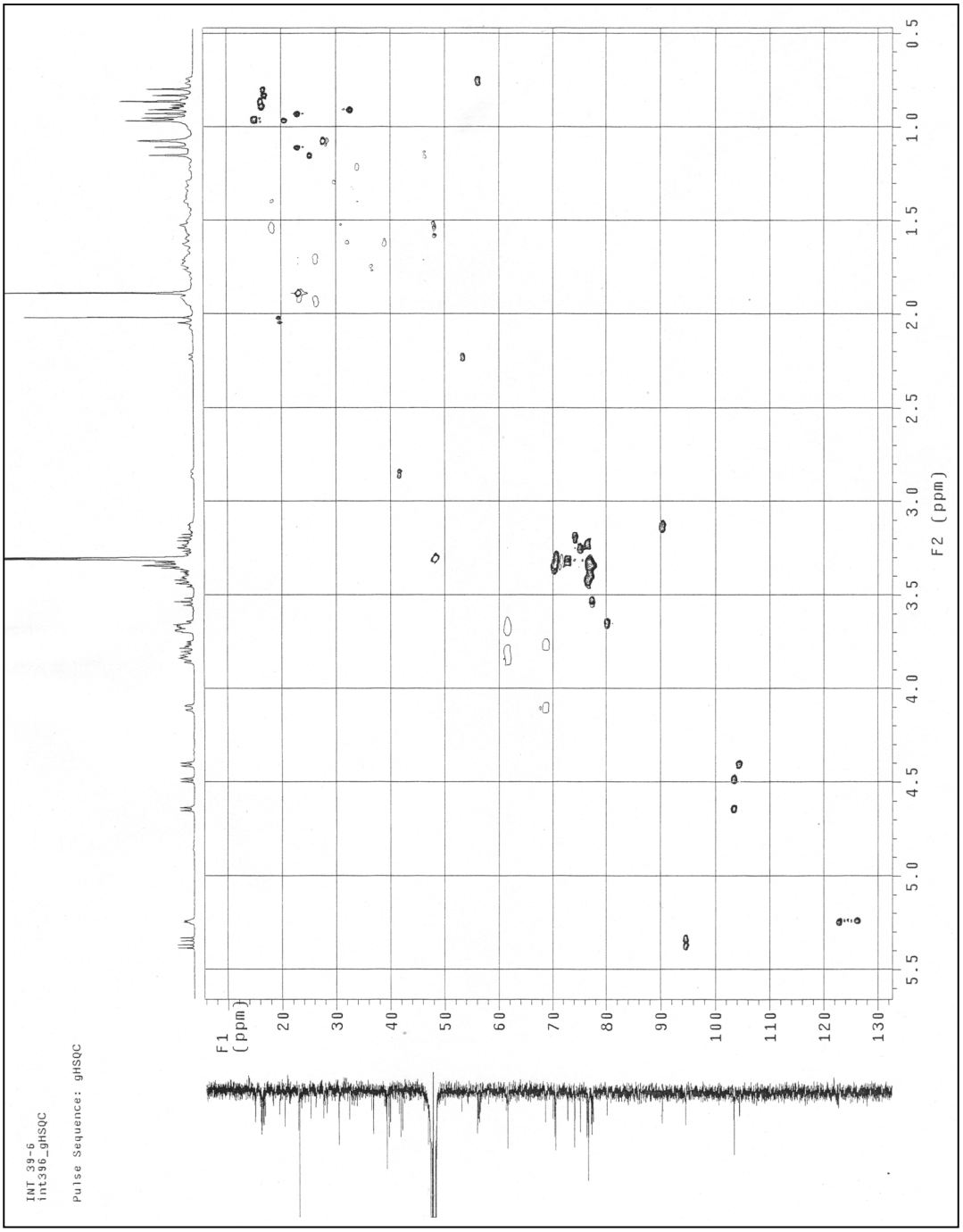


Figure B.15. HSQC spectra for *S. integrifolium* compound **4**.

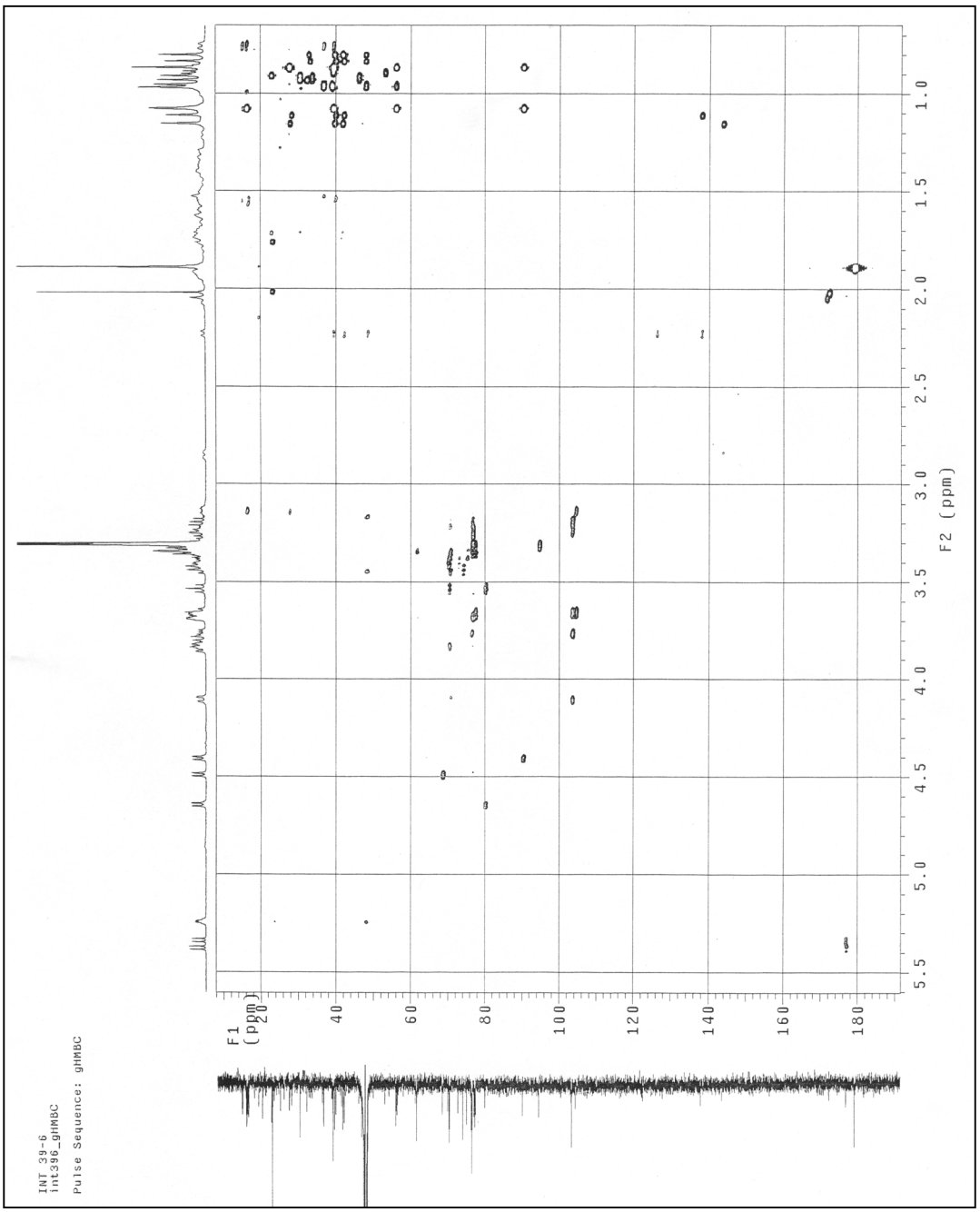


Figure B.16. HMBC spectra for *S. integrifolium* compound **4.**

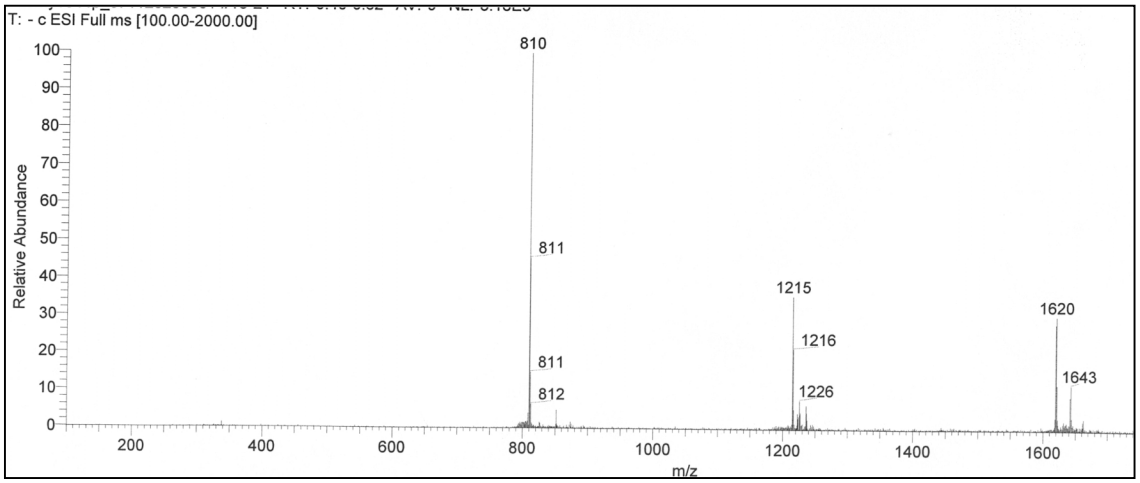


Figure B.17. ESI-MS negative mode spectra for compound **1** from *S. morhii.*

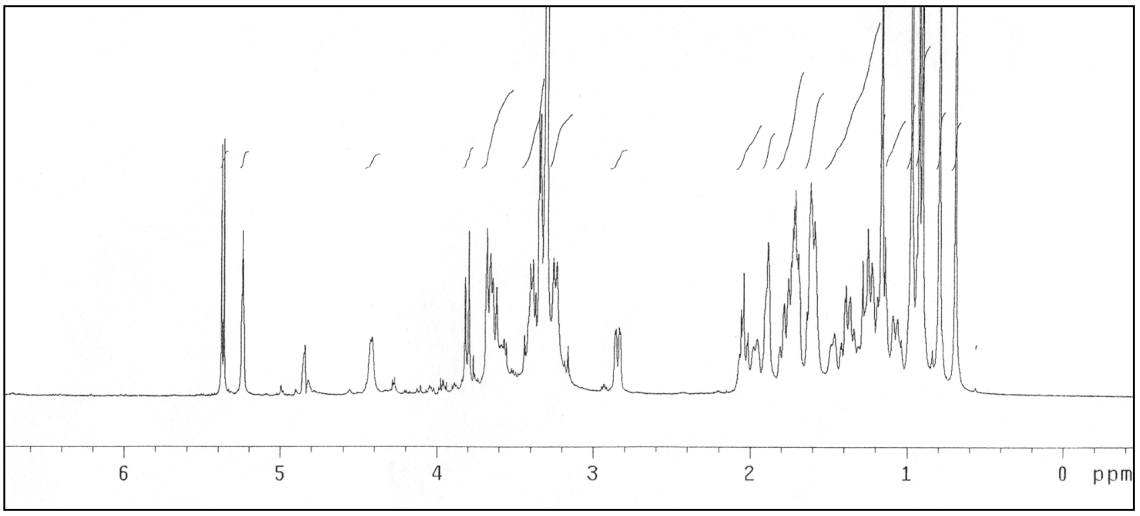


Figure B.18. ^{1}H-NMR spectra fro compound **1** from *S. morhii*.

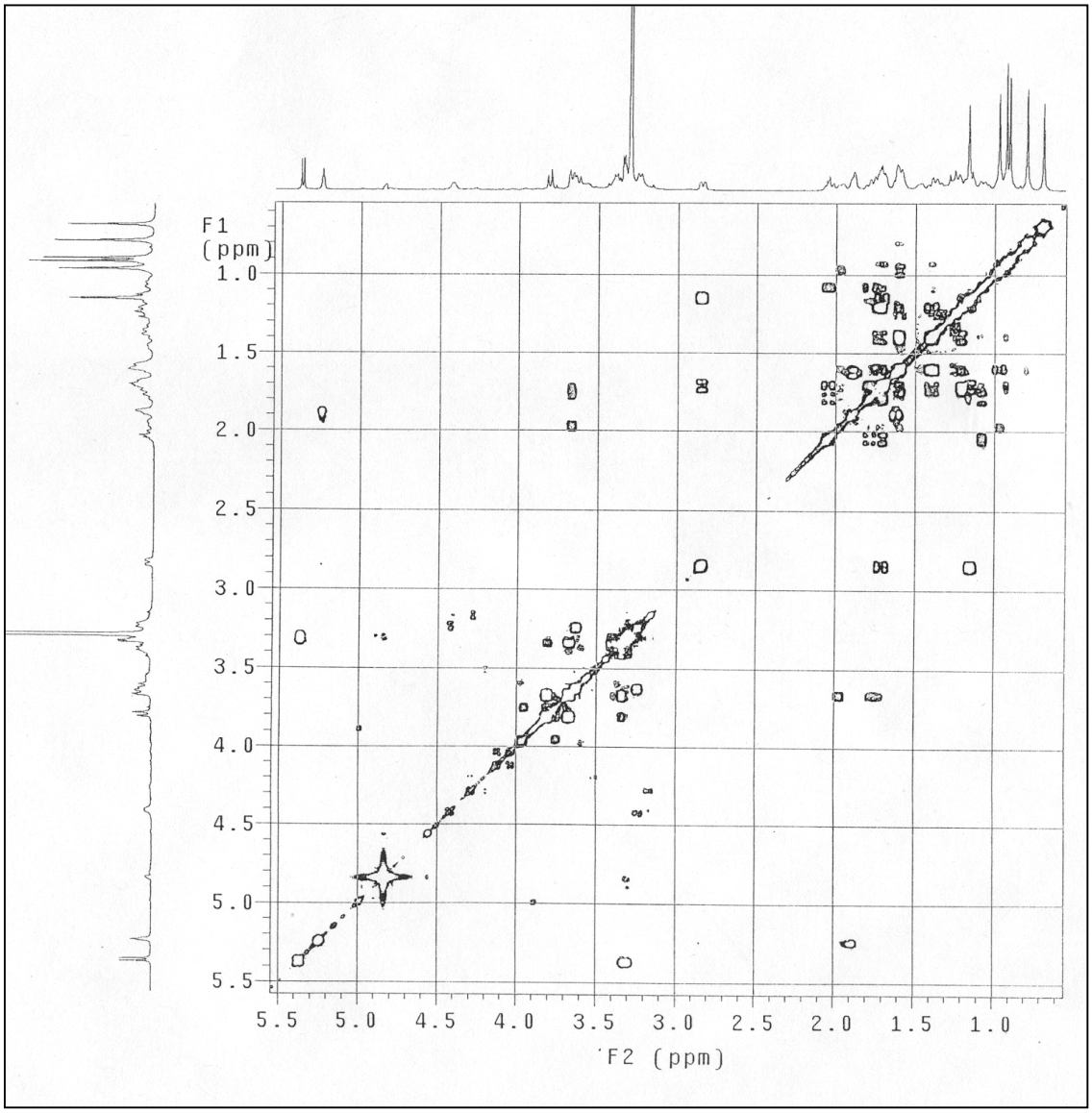


Figure B.19. COSY spectra for compound **1** from *S. morhii*.

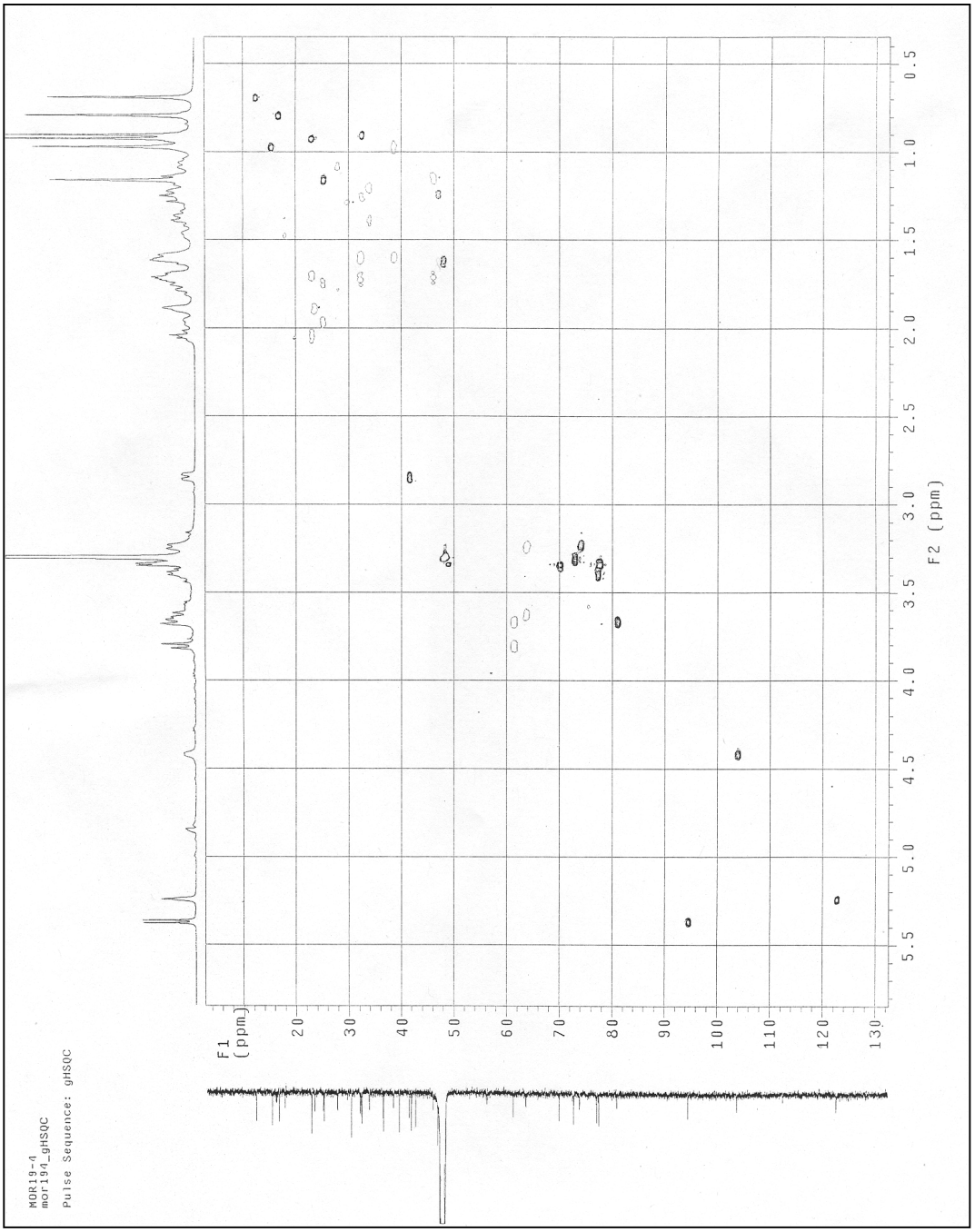


Figure B.20. HSQC spectra for compound **1** from *S. morhii*.

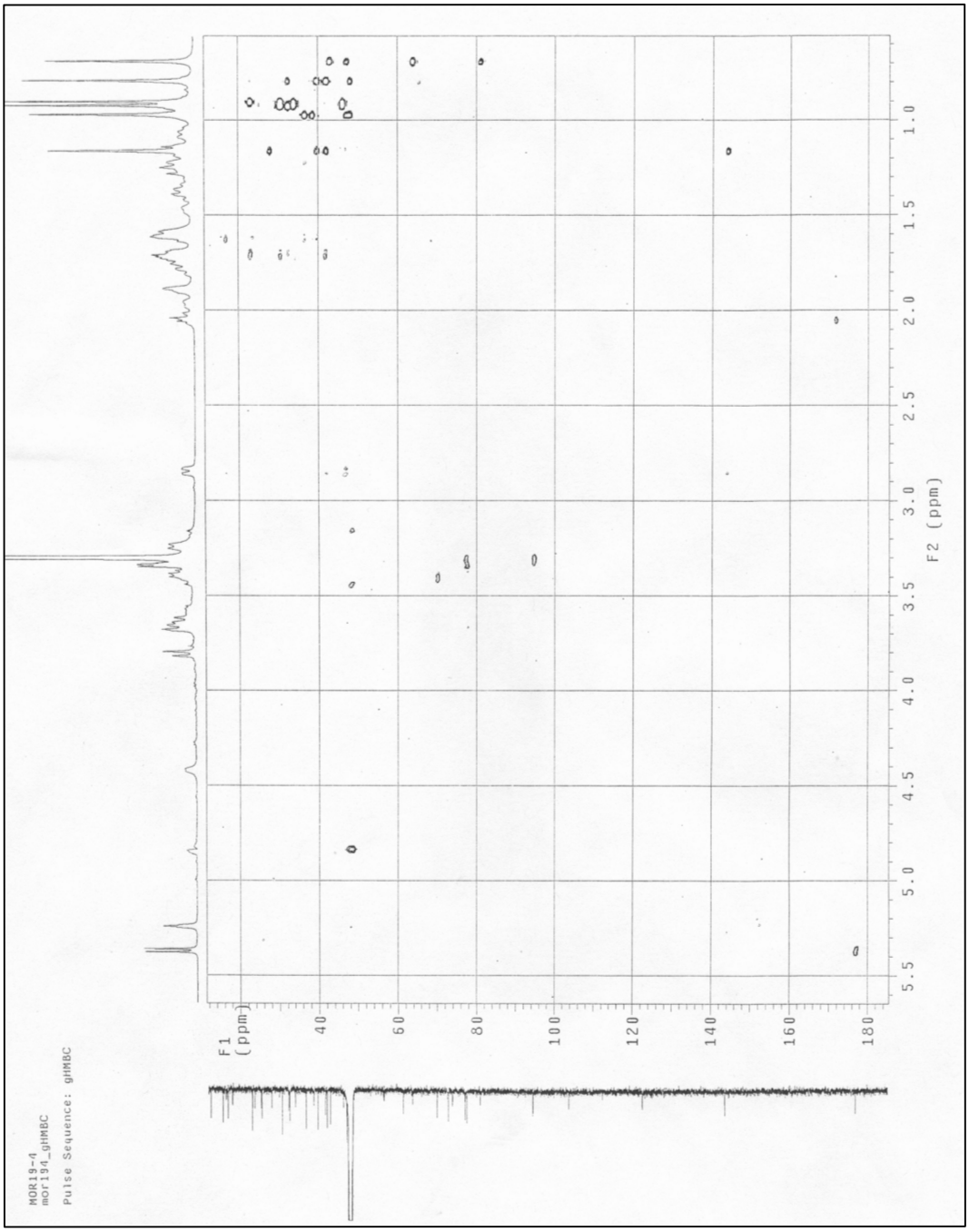


Figure B.21. HMBC spectra for compound **1** from *S. morhii*.

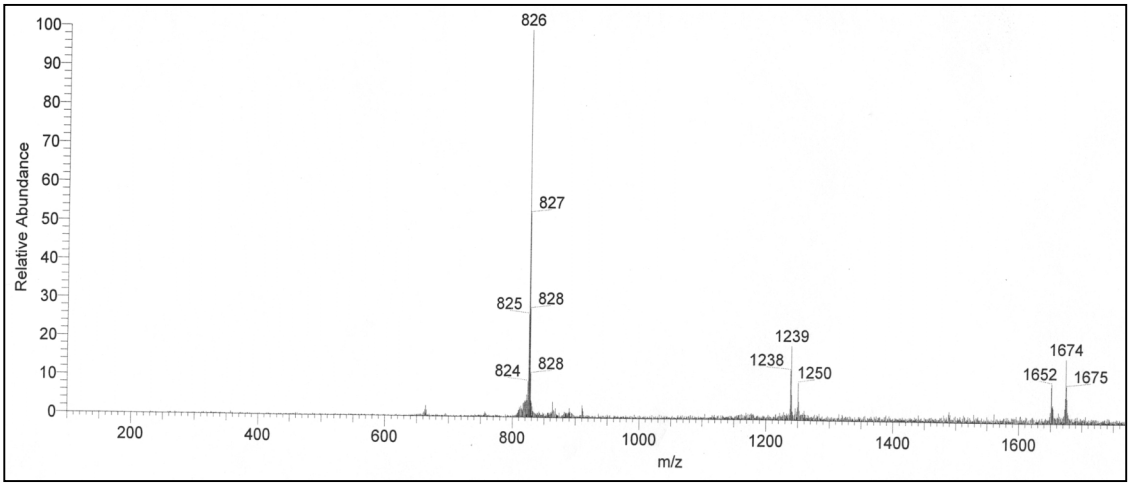


Figure B.22. ESI-MS negative mode spectra for *S. morhii* compound **2**.

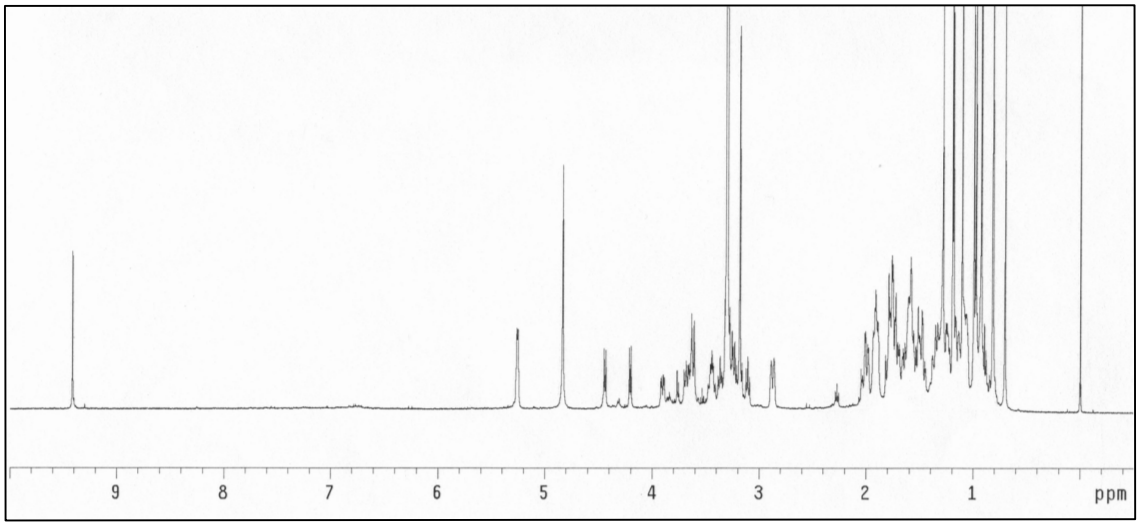


Figure B.23. ^{1}H-NMR spectra for *S. morhii* compound **2**.

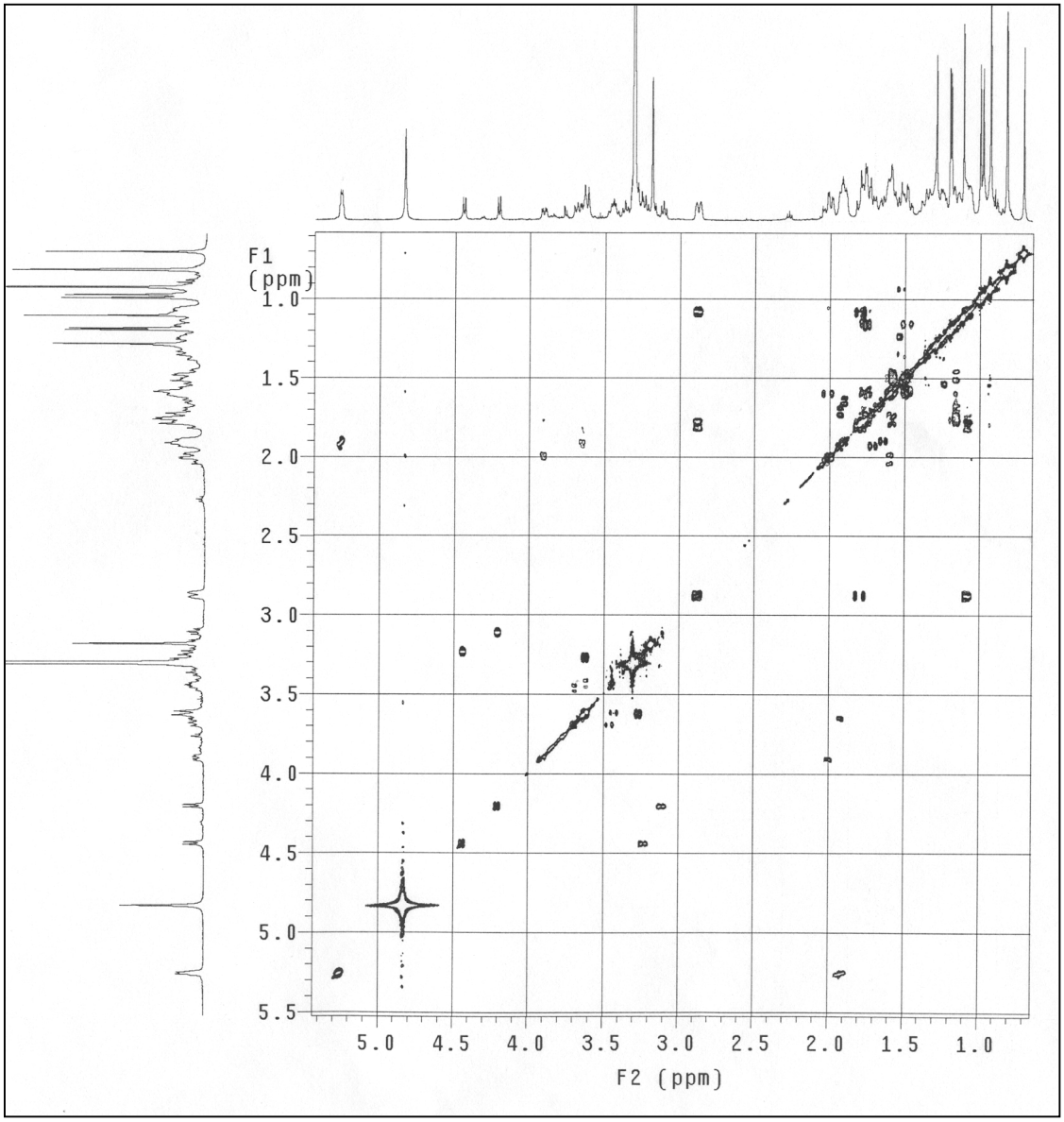


Figure B.24. COSY spectra of *S. morhii* compound **2**.

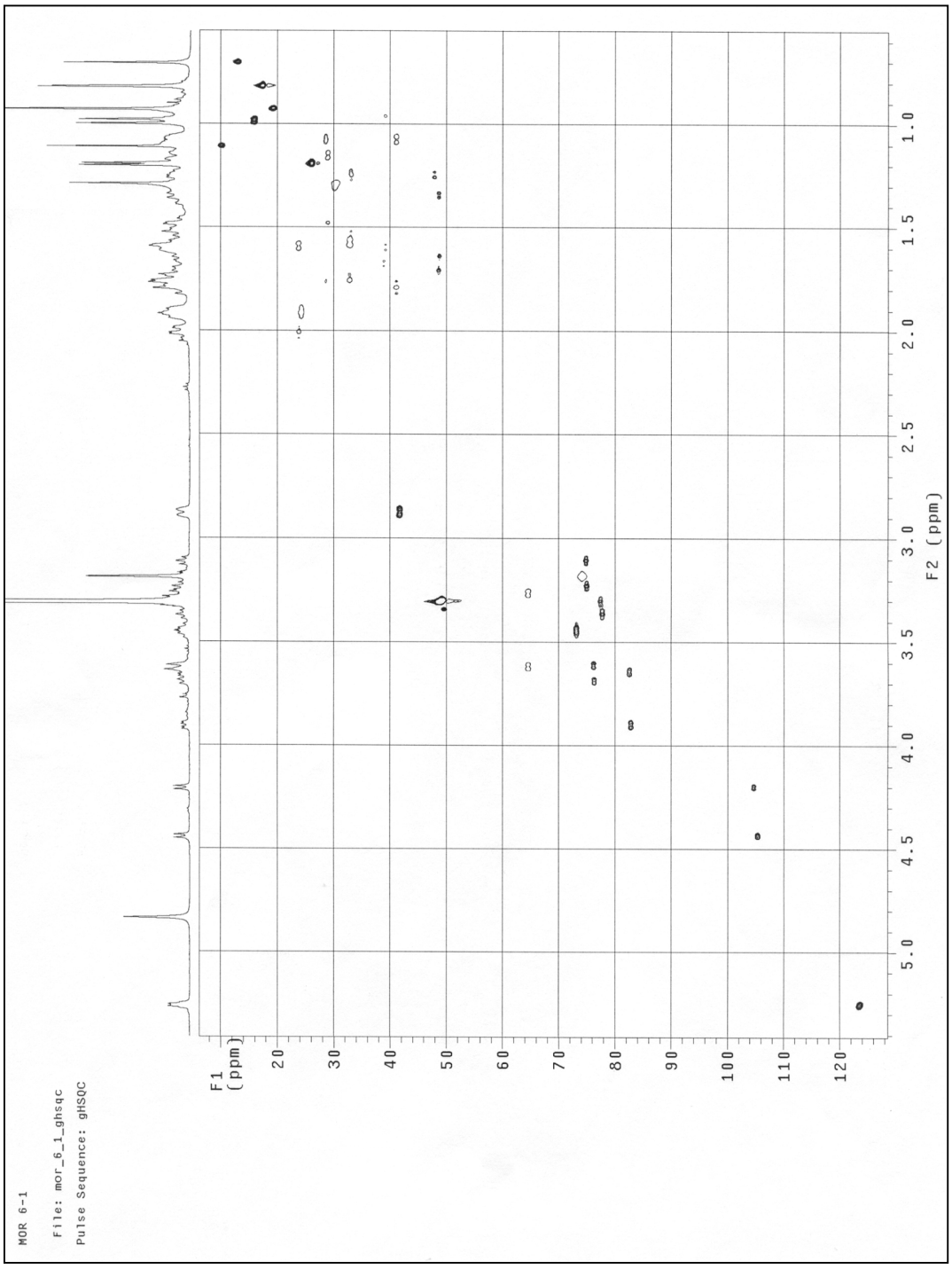


Figure B.25. HSQC spectra for *S. morhii* compound **2**.

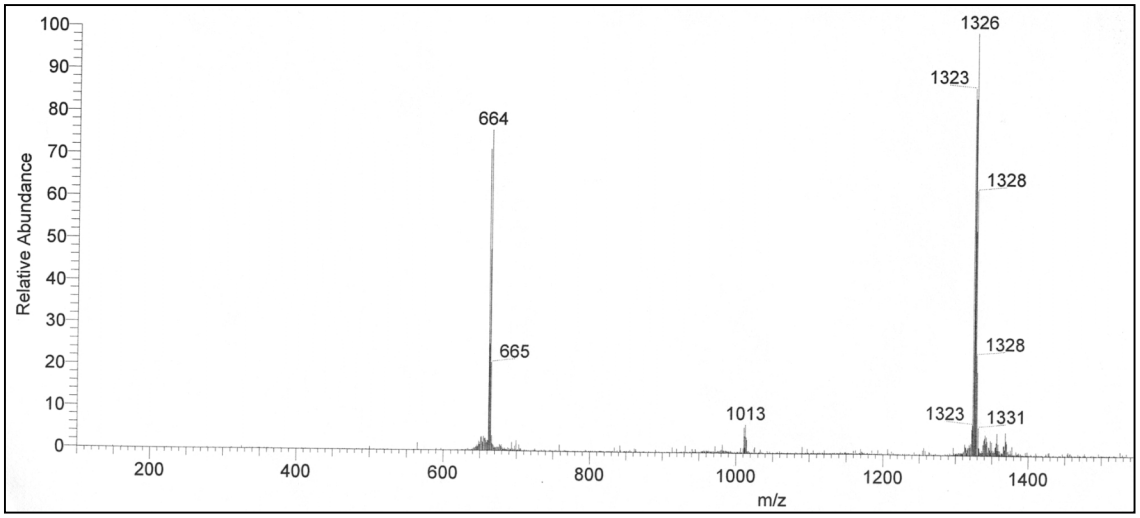


Figure B.26. ESI-MS negative mode spectra for *S. morhii* compound **3**.

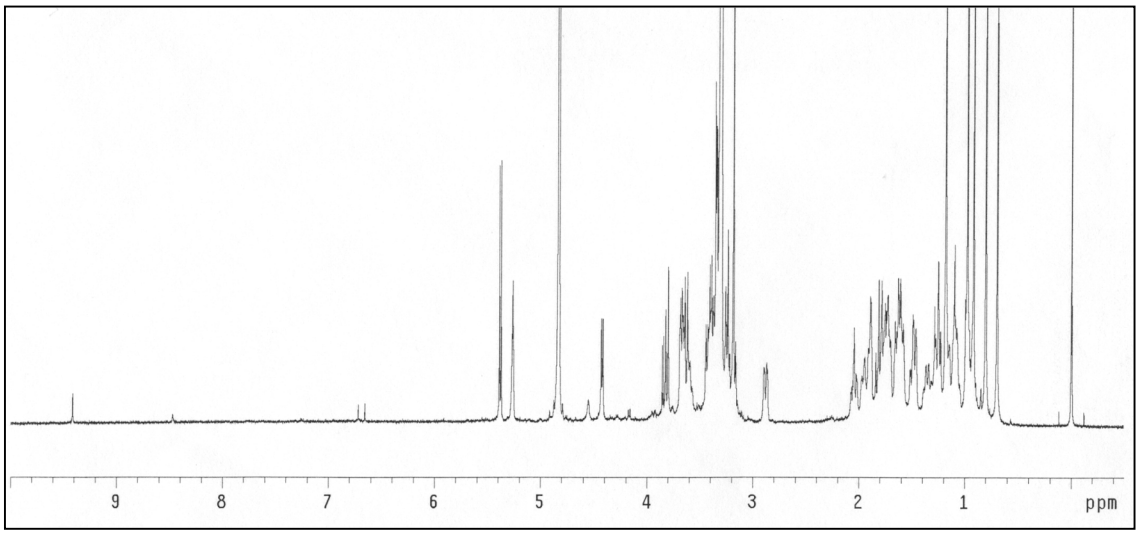


Figure B.27. ^{1}H-NMR spectra for *S. morhii* compound **3**.

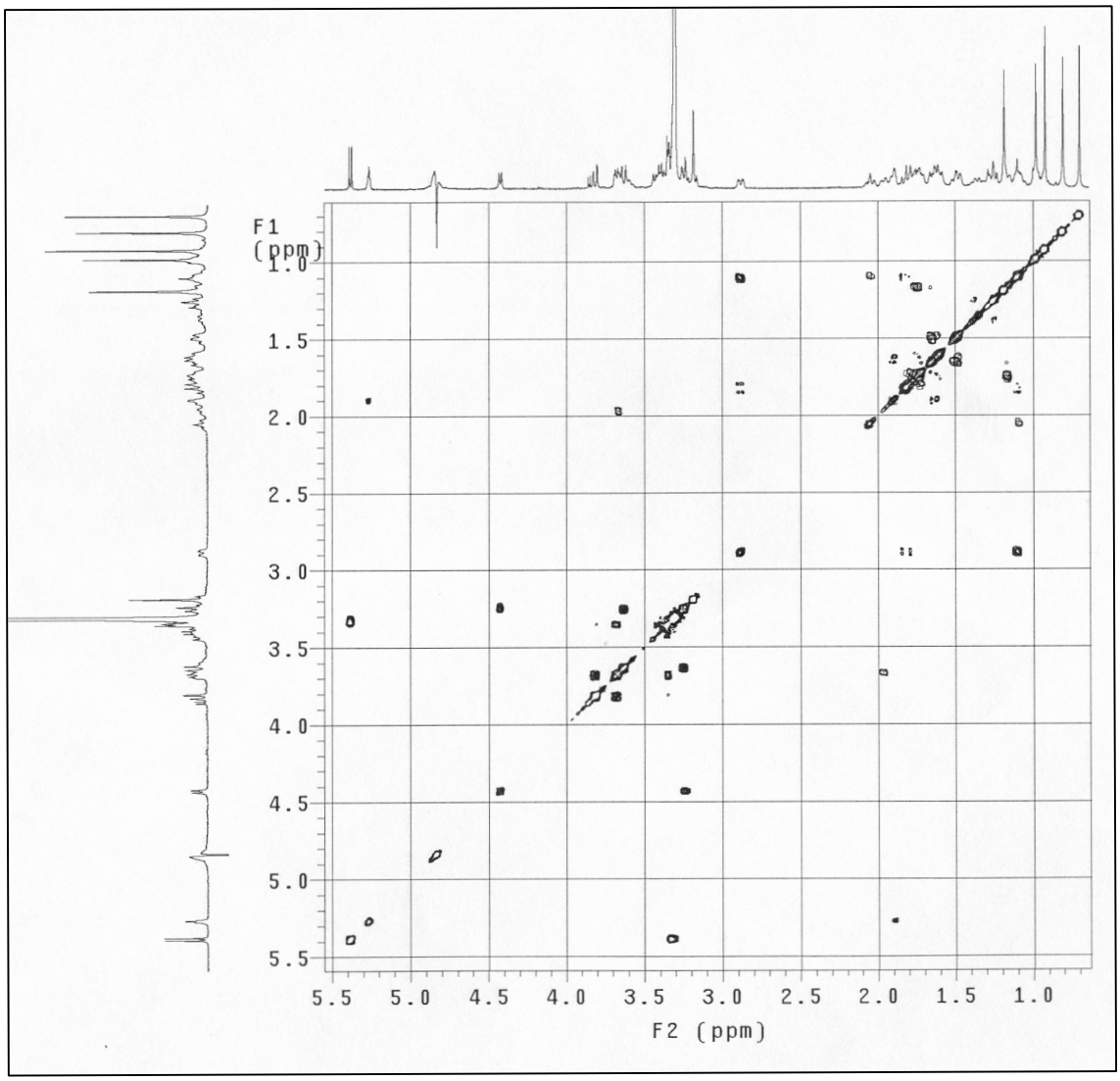


Figure B.28. COSY spectra for *S. morhii* compound **3**.

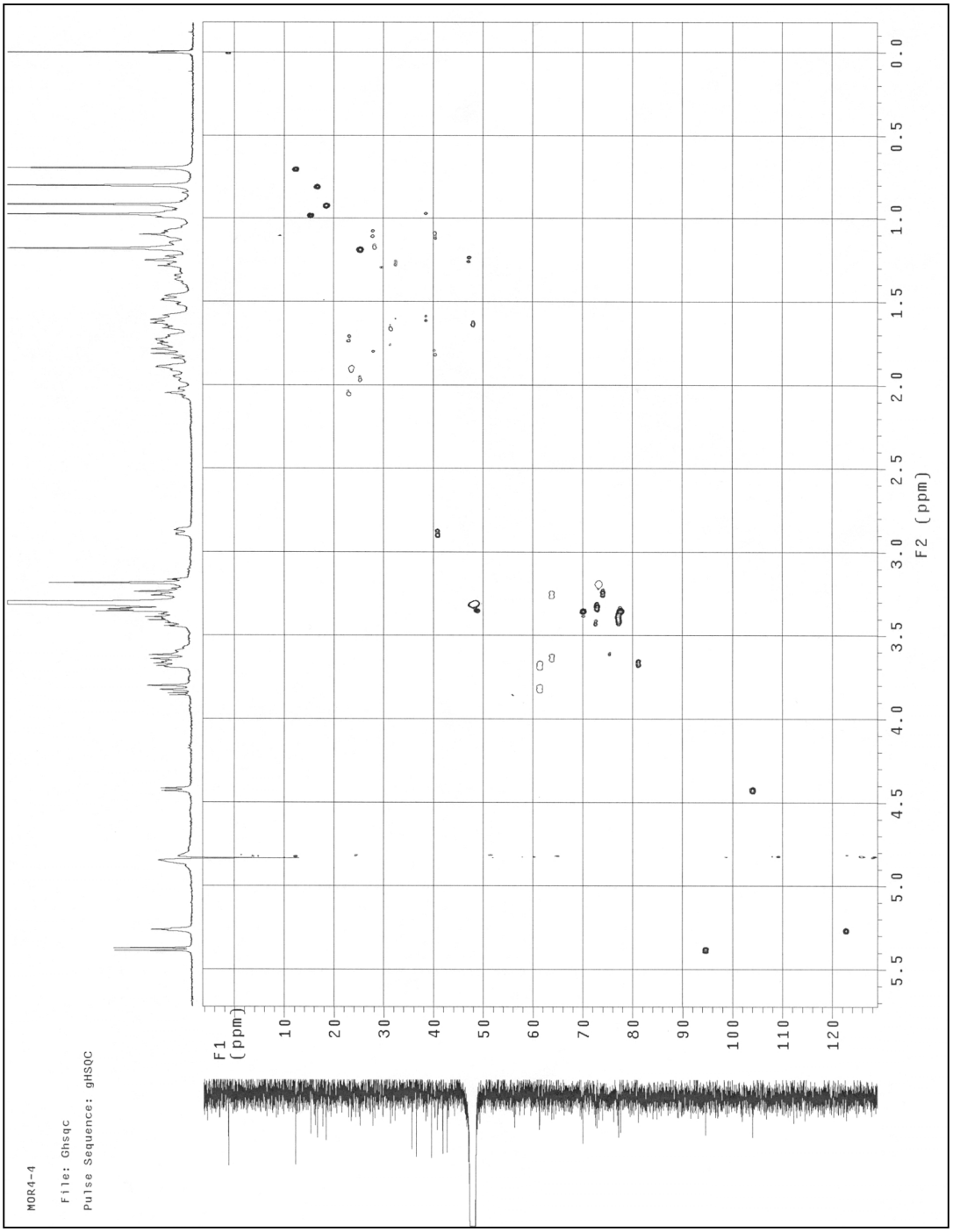


Figure B.29. HSQC spectra for *S. morhii* compound **3**.

Appendix C

	Rt	Saponin	Observed ESI/MS ions	COM	ALB	TER	LAC	WAS	PER	INT	BRA	MOR	AST	RAD	LIN
1	27.36-27.51		809.6 (100) [M-H]-, 647.6 (15) [M-hex-H]-, 1619.8 (4) [2M-H]-					x				x			
2	30.05-33.06	A	813.5 (100) [M-H]-, 651.3 (15) [M-hex-H]-, 489.4 (3) [M-hex-hex-H]-, 471.3 (7) [M-hex-hex-H2O-H]-, [M-agly]-							x	x		x	x	
3	30.53		811.7 (37) [M-H]-, 871.5 (100) [M+CH3COOH-H]-, 827.5 (58), 463.3 (27)							x					
4	30.55-35.35		941.6 (100) [M-H]-, 809.6 (7) [M-pen-H]-, 663.3 (2) [M-pen-hex-H]-		x			x	x	x					
5	31.84		975.5 (100) [M-H]-, 1950.4 (9) [2M-H]-, 814.4 (11) [M-Hex-H]-, 651.6 (8) [M-Hex-Hex-H]-, 489.3 (3) [M-Hex-Hex-Hex-H]-										x		
6	32.16		1223.6 (100) [M-H]-, 1091.6 (9) [M-pen-H]-, 929.6 (6) [M-pen-hex-H]-, 797.6 (8) [M-pen-hex-pen]-, 621.9 (3) [M-pen-hex-pen-GluA]-, 455.0 (6) [M-agly]-?	x											
7	32.24		795.5 () [M-H]-, 855.5 (100) [M+CH3COOH-H]-, 915.4 (43) [M+2CH3COOH-H]-, 633.5 (5) M-hex-H]-, 471.6 (4) [M-hex-hex-H]-							x					
8	32.44-32.54		825.6 (100) [M-H]-, 663.7 (8) [M-hex-H]-, 487.6 (1) [M-hex-GluA]-,					x				x			
9	33.59-36.93		971.6 (100) [M-H]-, 1943.8 (4) [2M-H]-					x		x					
10	33.67-34.37		811.6 (91) [M-H]-, 649.5 (20) [M-Hex-H]-, 601.8 (6), 473.5 (3) [M-hex-gluA-H]-, 871.4 (100) [M+CH3COOH-H]-, 1623.6 (13) [2M-H]-										x	x	
11	33.69		827.6 (100) [M-H]-, 1655.6 (2) [2M-H]-								x		x		
12	33.88		663.5 (100) [M-H]-, 487.5 (1) [M-GluA]-, 1327.6 (4) [2M-H]-									x			

Table C.1. Saponins (1-102) observed by negative-ion HPLC/ESI-MS in *Silphium* species and *Lindheimera texana.*

	Rt	Saponin	Observed ESI/MS ions	COM	ALB	TER	LAC	WAS	PER	INT	BRA	MOR	AST	RAD	LIN
13	34.38-36.98		809.7 (100) [M-H]-, 647.7 (6) [M-hex-H]-, 471.6 (1) [M-hex-gluA-H]-, 1619.7 (5) [2M-H]-				x	x	x	x	x	x	x		
14	34.87-37.7		971.6 (100) [M-H]-, 809.6 (51) [M-hex-H]-, 663.6 (16) [M-hex-hex-H]-					x		x					
15	34.8		797.6 (100) [M-H]-, 651.4 (7) [M-Rha-H]-, 471.1 (3), 453.3 (3)											x	
16	34.81		1013.5 (100) [M-H]-, 747.6 (85)		x										
17	35.09		1091.7 (100) [M-H]-, 797.5 (0.5) [M-pen-H]-,	x											
18	35.32-35.36		797.6 (100) [M-H]-, 857.5 (66) [M+CH3COOH-H]-, 1596.9 (4) [2M-H]-, 635.5 (26) [M-Hex-H], 473.5 (8) [M-Hex-Hex-H]-								x		x		
19	35.39-35.54		1091.6 (100) [M-H]-, 929.1 (11) [M-hex-H]-, 797.7 (13) [M-hex pen-H]-, 635.3 (5) [M-hex-pen-hex-H]-	x										x	
20	35.49		797.5 (100) [M-H]-, 857.4 (54) [M+CH3COOH-H]-, 1597.9 (2) [2M-H]-, 635.5 (38) [M-Hex-H], 473.5 (9) [M-Hex-Hex-H]-										x		
21	35.6		927.6 (100) [M-H-], 795.8 (3) [M-pen-H]-, 633.2 (1) [M-pen-hex-H]-, 471.7 (1) [M-pen-hex-hex-H]-							x					
22	35.61	B	649.4 (52) [M-H]-, 709.4 (100) [M-H+59]-, 1299.6 (33) [2M-H]-, 487.4 (9) [M-Glc-H]-, 559.5 (33) [M-90-H]-											x	
23	35.81		661.5 (100) [M-H]-, 485.6 (1) [M-H-GluA]-, 1323.5 (8) [2M-H]-									x			
24	36.07-40.51		925.6 (100) [M-H]-, 763.6 (10) [M-hex-H]]-, 1851.8 (6) [2M-H]-	x	x		x	x	x	x	x		x		
25	36.22		809.6 (100) [M-H]-, 647.6 (31) [M-hex-H]-, 485.5 (1) [M-hex-hex-H]-, 969.5 (19), 1619,7 (3) [2M-H]-					x							

Table C.1 *continued*. Saponins (1-102) observed by negative-ion HPLC/ESI-MS in *Silphium* species and *Lindheimera texana.*

	Rt	Saponin	Observed ESI/MS ions	COM	ALB	TER	LAC	WAS	PER	INT	BRA	MOR	AST	RAD	LIN
26	36.24-36-86		1075.7 (100) [M-H]-, 913.6 (6) [M-hex-H]-, 781.2 (2) [M-hex-pen-H]-	x										x	
27	36.63		955.6 (100) [M-H]-, 793.6 (13) [M-hex-H]-						x						
28	36.74	C	635.5 (28) [M-H]-, 1271.6 (45) [2M-H]-, 695.4 (100) [M-H+59]-, 473.3 (1) [M-Glc-H]-											x	
29	36.97		1013.4 [M-H]- (69), 909.4 (40), 873.4 (41), 747.6 (100), 597.1 (36)		x										
30	34.95-37.76	D	929.6 (100) [M-H]-, 1859 (5) [2M-H]-, 797.5 (7) [M-Ara-H]-, 779.7 (2) [M-Ara-H20-H]-, 635.6 (1) [M-Ara-Glc-H]-, 473 (0.4) [M-Ara-Glc-Glc-H]-							x				x	
31	37.5		1133.6 (100) [M+42-H]-?, 1091.6 (37), 1075.7 [M-H]-? (13), 943.6 (46) [M-pen-H]-, 929.6 (33), 797.8 (5), 617.6 (2), 487.5 (1)	x											
32	38.18	E	635.5 (34) [M-H]-, 1271.7 (47) [2M-H]-, 695.4 (100) [M-H+59]-, 473.5 (1) [M-Glc-H]-											x	
33	38.64		969.6 (100) [M-H]-											x	
34	36.07-38.91	F	913.7 (100) [M-H]-, 1827.8 (13) [2M-H]-, 781.7 (3) [M-Ara-H]-, 619.6 (1) [M-Ara-Glc-H]-, 455.3 (1) [M-Ara-Glc-Glc-H]-					x			x	x	x	x	
35	39.06		795.5 (100) [M-H]-, 855.3 (47) [M+CH3COOH-H]-, 1588.6 (1) [2M-H]-										x		
36	39.02-39.15		955.6 (100) [M-H]-, 793.5 (15) [M-hex-H]-, 617.7 (1) [M-hex-gluA-H]-, 569.6 (5), 455.7 (1) [M-hex-gluA-hex-H-]-		x		x	x			x				x
37	39.15		1117.6 (100) [M+42-H]-, 1075.6 (44) [M-H]-, 913.5 (20) [M-hex-H]-, 781.5 (5) [M-hex-pen-H]-	x											

Table C.1 *continued.* Saponins (1-102) observed by negative-ion HPLC/ESI-MS in *Silphium* species and *Lindheimera texana.*

	Rt	Saponin	Observed ESI/MS ions	COM	ALB	TER	LAC	WAS	PER	INT	BRA	MOR	AST	RAD	LIN
38	39.19		797.6 (100) [M-H]-, 1595.8 (4) [2M-H]-, 857.5 (37) [M+CH3COOH-H]-, 635.5 (6) [M-hex-H]-, 473.6 (3) [M-hex-hex-H]-							x					
39	39.26		807.6 (100) [M-H]-, 645.5 (13) [M-hex-H]-, 469.6 (1) [M-hex-GluA-H]-, 1615.6 (2) [2M-H]-									x			
40	39.57		797.6 (100) [M-H]-, 857.5 (25) [M+CH3COOH-H]-, 635.4 (4) [M-hex-H]-, 473.5 (3) [M-hex-hex-H]-							x					
41	39.8		959.6 (100) [M-H]-, 797.6 (16) [M-hex-H]-, 635.5 (1)[M-hex-hex-H]-											x	
42	40.29		811.6 (100) [M-H]-, 635.7 (3) [M-gluA-H]-, 1623.8 (3) [2M-H]-							x					
43	40.37		957.6 (100) [M-H]-, 1913.6 (3) [2M-H]-, 795.5 (4) [M-1hex-H]-, 649.6 (6) [M-hex-146-H]-											x	
44	40.47		731.6 (100) [M-H]-, 569.5 (7) [M-hex-H]-, 997.4 (46), 893.4 (13), 835.5 (6) [997.4-hex-H]-?												x
45	40.75-41.02		647.5 (100) [M-H]-, 1295.6 (16) [2M-H]-, 471.6 (1) [M-GluA-H]-					x				x			
46	40.88-41.41		943.6 (100) [M-H]-, 781.5 (7) [M-hex-H]-, 955.7 (49)	x										x	
47	40.99-43.51	O	793.6 (100) [M-H]-, 1587.9 (6) [2M-H]-, 631.7 (8) [M-hex-H]-, 455.7 (1) [M-hex-gluA-H]-				x	x	x		x	x	x		x
48	41.01-41.52		779.5 (100) [M-H]-, 731.7 (17), 617.7 (2) [M-hex-H]-, 455.6 (2) [M-hex-hex-H]-,							x					x
49	41.84-41.89		779.5 (30) [M-H]-, 821.5 (100), 617.5 (3) [M-hex-H]-, 455.5 (3) [M-hex-hex-H]-							x					x
50	42.18		985.5 (17) [M-H]-, 1970.9 (1) [2M-H]-, 797.6 (100), 635.6 (7), 473.8 (2), 455.5 (2)											x	

Table C.1 *continued.* Saponins (1-102) observed by negative-ion HPLC/ESI-MS in *Silphium* species and *Lindheimera texana.*

	Rt	Saponin	Observed ESI/MS ions	COM	ALB	TER	LAC	WAS	PER	INT	BRA	MOR	AST	RAD	LIN
51	42.39		777.7 (21) [M-H]-, 837.5 (100) [M+CH3COOH-H]-, 615.2 (3) [M-hex-H]-,							x					
52	42.8-42.85		731.6 (100) [M-H]-, 569.5 (12) [M-hex-H]-, 997.4 (46), 893.4 (35)		x										x
53	42.84		781.7 (100) [M-H]-, 943.3 (16), 979.4 (27)	x											
54	42.85-42.96	G	811.6 (100) [M-H]-, 1623.6 (2) [2M-H]-,										x	x	
55	43.11		835.5 (100) [M-H]-, 673.5 (6) [M-hex-H]-, 569.6 (15),						x						
56	43.21-44.65	K	779.6 (38) [M-H]-, 839.5 (100) [M+CH3COOH-H]-, 617.8 (5) [M-hex-H]-, 455.5 (1) [M-hex-hex-H]-						x	x			x	x	
57	43.46	H	795.7 (100) [M-H]-, 1591.8 (9) [2M-H]-, 633.4 (3) [M-Glc-H]-											x	
58	43.77		663.5 (100) [M-H]-, 487.5 (1) [M-GluA]-, 1327.6 (9) [2M-H]-									x			
59	44.09		853.6 (100) [M-H]-, 1707.6 (6) [2M-H]-							x					
60	44.12--44.43		809.6 (100) [M-H]-, 663.6 (14) [M-146-H]-, 487.4 (1) [M-146-gluA-H]-, 471.3 (1) [M-146-gluA-OH-H]-		x		x	x			x				
61	44.22		663.5 (100) [M-H]-, 487.4 (3) [M-GluA]-, 1327.5 (4) [2M-H]-									x			
62	44.4		765.6 (100) [M-H]-, 823.5 (31)	x											
63	44.5		881.5 (100) [M-H]-, 1763.5 (2) [2M-H]-							x					
64	44.59-45.99		779.3 (48) [M-H]-, 839.4 (100) [M+CH3COOH-H]-, 617.7 (100) [M-hex-H]-, 455.0 (3) [M-hex-hex-H]-										x	x	x

Table C.1 *continued*. Saponins (1-102) observed by negative-ion HPLC/ESI-MS in *Silphium* species and *Lindheimera texana.*

	Rt	Saponin	Observed ESI/MS ions	COM	ALB	TER	LAC	WAS	PER	INT	BRA	MOR	AST	RAD	LIN
65	45.16		775.5 (20) [M-H]-, 835.5 (100) [M+CH3COOH-H]-, 613.4 (12) [M-hex-H]-, 569.6 (86)												x
66	45.17-47.85	M	941.7 (100) [M-H]-, 1883.7 (1) [2M-H]-, 779.6 (9) [M-Hex-H]-, 749.6 (10)					x	x	x	x				
67	45.5		775.5 (16) [M-H]-, 835.5 (100) [M+CH3COOH-H]-, 613.4 (15) [M-hex-H]-, 569.6 (79)												x
68	45.65		823.6 (100) [M-hex-H]-, 1021.4 (18) [M-H]-, 985.3 (10) [M-36-H]-	x											
69	46.04		649.5 (100) [M-H]-, 487.4 (11) [M-hex-H]-						x						
70	46.14		835.4 (100) [M-H]-, 659.7 (14) [M-gluA-H]-, 569.6 (34), 483.6 (2) [M-gluA-gluA-H]-												x
71	46.22		661.5 (100) [M-H]-, 485.4 (2) [M-GluA-H]-, 1323.6 (4) [2M-H]-									x			
72	46.29		879.5 (100) [M-H]-, 747.6 (18) [M-pen-H]-, 615 (2) [M-pen-pen-H]-							x					
73	46.39-46.51		779.5 (100) [M-H]-, 1559.8 (11) [2M-H]-, 629.5 (19), 585.5 (5) 471.4 (1)		x	x	x				x				
74	46.49		821.5 (100) [M-H]-, 659.6 (4) [M-hex-H]-, 779.5 (7)												x
75	46.62		853.5 (100), 709.4 (17), 649.2 (11), 487.4 (10)										x		
76	46.78		955.6 (100) [M-H]-, 835.6 (60), 793.6 (4) [M-hex-H]-, 631.6 (6) [M-hex-hex-H]-						x						
77	46.78		821.5 (100) [M-H]-, 659.6 (4) [M-hex-H]-, 779.5 (21)												x
78	46.98		809.6 (100) [M-H]-, 1619.8 (1) [2M-H]-					x							

Table C.1 *continued.* Saponins (1-102) observed by negative-ion HPLC/ESI-MS in *Silphium* species and *Lindheimera texana.*

	Rt	Saponin	Observed ESI/MS ions	COM	ALB	TER	LAC	WAS	PER	INT	BRA	MOR	AST	RAD	LIN
79	47.06		865.6 (100) [M-H]-, 823.7 (57), 661.4 (5), 457.6 (4)							x					
80	47.67		879.5 (100) [M-H]-							x					
81	47.74-49.09		835.5 (100) [M-H]-, 673.5 (4) [M-hex-H]-, 647.5 (51), 569.7 (19)						x						x
82	48.07		851.4 (100) [M-H]-, 1703.9 (2) [2M-H]-, 747.6 (69)		x										
83	48.17		645.5 (100) [M-H]-, 469.6 (4) [M-GluA-H]-,1291.6 (3) [2M-H]-									x			
84	48.19		865.5 (100) [M-H]-							x					
85	48.2		781.6 (100) [M-H]-, 649.6 (30) [M-pen-H]-, 847.4 (22), 803.5 (27)	x											
86	49.16		807.5 (100) [M-H]-, 1615.4 (1) [2M-H]-, 661.6 (2) [M-146-H]-					x							
87	50.08-50.26		647.5 (100) [M-H]-, 1295.6 (16) [2M-H]-, 471.5 (3) [M-GluA-H]-	x	x	x	x	x	x	x	x	x			
88	50.91		765.5 (100) [M-22-H]-, 787.5 (47) [M-H]- 655.5 (51) [M-pen-H]-	x											
89	51.24		777.6 (100) [M-H]-, 645.5 (28) [M-pen-H]-									x			
90	52.49-52.56		647.5 (100) [M-H]-, 1295.6 (13) [2M-H]-, 471.5 (1) [M-GluA-H]-				x	x				x	x		
91	52.49-54.07		763.6 (100) [M-H]-, 1527.8 (40) [2M-H]-, 823.6 (100) [M+CH3COOH-H]-	x					x						
92	53.17-53.53	I	781.7 (100) [M-H]-, 1564.6 (20) [2M-H]-, 617.1 (1) [M-Glc-H]-, 455.3 (1) [M-Glc-Gal-H]-											x	
93	53.32-53.39		793.6 (100) [M-H]-, 647.7 (7) [M-146-H]-, 1587.7 (2) [2M-H]-		x			x			x				x
94	54.6		617.5 (100) [M-H]-, 1235.6 (5) [2M-H]-, 455.5 (1) [Agly-H]- or [M-Hex-H]-								x				

Table C.1 *continued.* Saponins (1-102) observed by negative-ion HPLC/ESI-MS in *Silphium* species and *Lindheimera texana.*

	Rt	Saponin	Observed ESI/MS ions	COM	ALB	TER	LAC	WAS	PER	INT	BRA	MOR	AST	RAD	LIN
95	54.76		779.4 (100) [M-H]-, 617.4 (4) [M-hex-H]-, 455.4 (5) [M-hex-hex-H]-												x
96	55.21-55.24		731.6 (100) [M-H]-, 835.2 (34), 569.3 (10) [M-hex-H]-		x										x
97	55.85		763.6 (100) [M-H]-, 1527.8 (5) [2M-H]-								x				
98	55.87		645.5 (100) [M-H]-, 1291.6 (6) [2M-H]-, 469.7 (2) [M-gluA-H]-					x				x			
99	57.02		731.6 (100) [M-H]-, 835.2 (84), 569.3 (30) [M-hex-H]-												x
100	57.16-57.51		763.5 (100) [M-H]-, 617.6 (9) [M-Rha-H]-		x			x			x				
101	58.69		763.6 (100) [M-H]-, 1528.8 (2) [2M-H]-								x				
102	59.02-60.51	N	631.5 (100) [M-H]-, 1263.7 (16) [2M-H]-, 455.6 (1) [M-gluA-H]-	x	x	x	x	x	x	x	x	x	x	x	x

Table C.1 *continued*. Saponins (1-102) observed by negative-ion HPLC/ESI-MS in *Silphium* species and *Lindheimera texana.*

Species abbreviations:
COM= *S. compositum*
ALB= *S. albiflorum*
TER= *S. terebinthinaceum*
WAS= *S. wasiontense*
PER= *S. perfoliatum*
INT= *S. integrifolium*
BRA= *S. brachiatum*
MOR= *S. morhii*
AST= *S. asteriscus*
RAD= *S. radula*
LIN= *Lindheimera texana*

Appendix D: Tom J. Mabry's Natural Products Chemistry Program: 1960-2007[7]

Abstract

This paper presents an overview of Dr. Mabry's accomplishments in his career as a natural product chemist, first at the University of Zürich as a post-doctorate and since 1962 as a faculty member at the University of Texas at Austin in the Department of Botany until the late 1990's when the Biological Sciences programs at UT-Austin were completely reorganized; then until his retirement in 2006, he was a member of the Molecular Cell and Developmental Biology faculty.

[7] This paper was originally published as: "Tom J. Mabry's Natural Products Chemistry Program: 1960-2007" Lalita M. Calabria with Tom J. Mabry. 2007. Natural Product Communications 2, 959-968.

Introduction

This issue of Natural Product Communications is dedicated to the 75th birthday of Tom J. Mabry, Professor Emeritus at the University of Texas at Austin. For this paper, I outline Dr. Mabry's outstanding achievements to the field of natural products chemistry, and I add Dr. Mabry as second author in appreciation of his many comments and contributions. As his last Ph.D. student, it is my special honor to describe the personal and professional impact Dr. Mabry has made on my life, a sentiment that is shared by many, if not all of his students, post-doctorates and colleagues over the last 40 years as a Professor at UT-Austin. To prepare this report of his accomplishments I pursued many of his publications (nearly 700!), including 15 books and several dozen chapters, as well as some of his students' dissertations and theses. In addition to reviewing these written works, I also summarize Dr. Mabry's accounts of several of his major projects, beginning with his Ph.D. studies on the coenzymatic activity of Vitamin C and his post-doctoral investigations marking the discovery of a new class of plant pigments found initially in beets, the betalains, which proved significant for the plant sciences and thus for his career at UT-Austin. In addition, I interviewed a few of his more than 70 MS and Ph.D. students and about ten of his hundreds of post doctorates, as well as several of his faculty colleagues.

Although Dr. Mabry modestly credits his career success to "the luck of pursuing the right projects with the right people at the right time", when I began to review his many accomplishments, awards and memberships, it was clear that most of his success came from his own intense efforts. Dr. Mabry was instrumental in organizing the Phytochemical Society of North America in 1966 and early in his career, Prof. T.W. Goodwin, a distinguished biochemist in England, praised Mabry as the "Father of Modern Phytochemistry in the United States". Mabry has received numerous awards including the Alexander von Humboldt Senior Scientist Award for research with Prof. Dietmar Behnke at the University of Heidelberg, Germany; American Chemical Society Award for the "Application of Chemistry to Food and Agriculture;" "Pergamon Phytochemistry Prize;" UT-Austin Graduate School's "Outstanding Doctoral Teaching

Award;" and American Society of Pharmacognosy's "Research Achievement Award". It would take much more than this paper to describe all of Dr. Mabry's many contributions to natural products chemistry, thus, I will provide just a few examples of the scope and breadth of his work while highlighting his extraordinary abilities as a teacher and scientist.

Dr. Mabry's 1956-1960 Ph.D. studies expanded on an earlier finding that L-ascorbic acid (Vitamin C) had unusually high coenzymatic properties for the hydrolysis of mustard oil glucosides, while D-ascorbic acid, which has no Vitamin C activity, showed no coenzymatic activity for the hydrolysis of these same glycosides. Since many analogs of L-ascorbic acid had been tested in the 1930's in guinea pigs and were found to exhibit low or no Vitamin C activity, Mabry synthesized the same analogs and established that they each had low or no activity for the hydrolysis of the glycosides, that is, the same pattern they had exhibited for Vitamin C activity (Figure 1). His findings suggest that the Vitamin C activity of L-ascorbic acid is related to its coenzymatic properties and not to its well-established antioxidant potential [see discussions of these findings in 1, 2]. Dr. Mabry strongly encourages further investigations to determine which reactions in humans, if any, involve L-ascorbate as a coenzyme in its role as Vitamin C. (Mabry suspects the coenzyme activities of L-ascorbic acid could be involved with the synthesis of new tissue to replace that degraded by colds, aging and, in earlier times, in Vitamin C deficiency diseases such as scurvy).

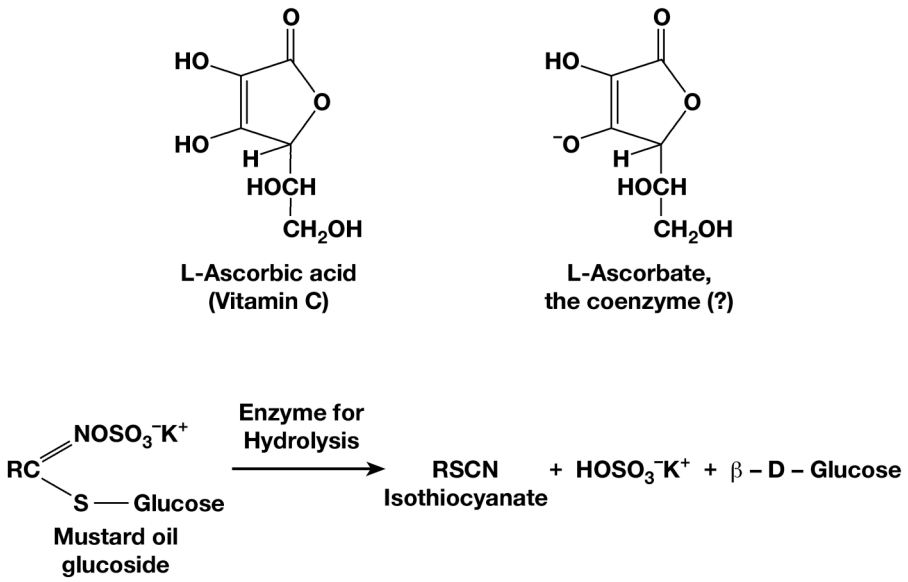


Figure 1. The rate of hydrolysis of mustard oil glucosides is markedly increased by L-ascorbic acid (Vitamin C), but not by D-ascorbic acid, which has no vitamin activity. Synthetic analogs of L-ascorbic acid exhibited comparable patterns of Vitamin C and coenzymatic activities.

After receiving his Ph.D. in Chemistry in 1960 from Rice University, he joined the group of Professor Andre Dreiding (a specialist for reaction mechanisms) in the Organic Chemistry Institute, University of Zürich, Switzerland. He was assigned to tackle Dreiding's only plant project, namely to determine the structure of the red beet pigment, betanin, that was thought to be a member of the well-known and widespread pigments, the anthocyanins (roses, red wine, etc.); however, it was referred to as a "nitrogenous anthocyanin" because both of the breakdown products from betanin contained nitrogen. Dr. Mabry recognized that previous investigators in Dreiding's labs had used rather harsh conditions that degraded the beet pigment to the nitrogenous breakdown products. Therefore he employed very mild procedures for derivatizing betanidin, the aglycone of betanin: for example, methylation with diazomethane in ether at neutral pH and room temperature. Using variations of such mild procedures Mabry was able to isolate several derivatives of betanidin that contained all the carbon, oxygen and nitrogen atoms found in this aglycone of betanin; he named these derivatives neobetanidins. In collaboration with Hugo Wyler and others in Dreiding's group, Mabry

quickly established the structures of the neobetanidins, as well as betanidin and betanin [3, 4]. These findings startled the plant world because the new structures were not related to those for anthocyanins; thus, a new class of pigments, the "betalains" was recognized (Figure 2) [the name was selected by Mabry and Dreiding and published by them in 1968 [5].

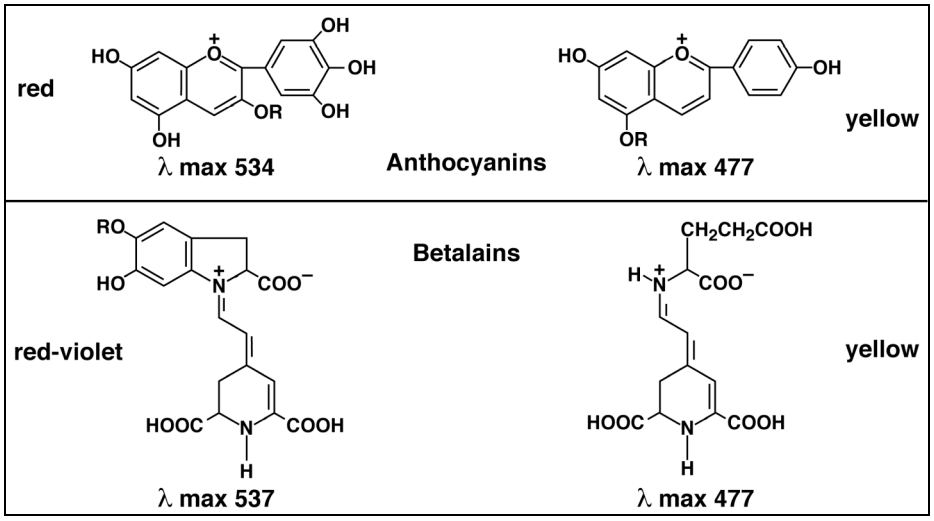


Figure 2. In the late 1960's, Mabry and Dreiding recognized the betalains as a new class of pigments that are found in nine plant families; these families produce none of the well-known and widely distributed plant pigments, the anthocyanins.

In 1962, Dr. Mabry accepted a position in the plant biology program at the University of Texas at Austin to organize a major natural products chemistry program in the Department of Botany. Professor B. L. Turner (now Emeritus), a distinguished plant systematist at UT-Austin, and his plant physiologist colleague, the late Dr. Ralph Alston, were instrumental in recruiting Dr. Mabry from the University of Zürich (see Figure 3). With the support of Turner and Alston, Mabry assembled a group of truly outstanding chemistry and biochemistry post-doctorates and many bright and motivated botany graduate students who wanted to learn all they could about chemical techniques for studying plants (Figure 4). Moreover, he was able to obtain funding for the new research programs from NSF, NIH, USDA and many foundations, including especially the Robert

A. Welch Foundation in Houston, which still today supports his program; he also obtained special funds for NMR, MS, UV and chromatographic equipment.

Prior to 1962, botanists at UT-Austin had observed that when extracts of *Baptisia* (family Leguminosae) and several other genera were analyzed by two-dimensional paper chromatography and then examined under UV light, the patterns of brightly colored flavonoid "spots" were often reliable markers for species identifications and for recognizing hybrids (Figure 3) [6, 7]. Since it was clear that having actual chemical structures for each "spot" would enormously increase the reliability of using these data for plant systematics, Mabry's initial goal was to structurally characterize the flavonoids in *Baptisia* species. Thus, a leading biochemical systematic research program was established in Mabry's lab, one that would continue to produce many new natural products for systematics, evolutionary studies and medicine for the next 40 years.

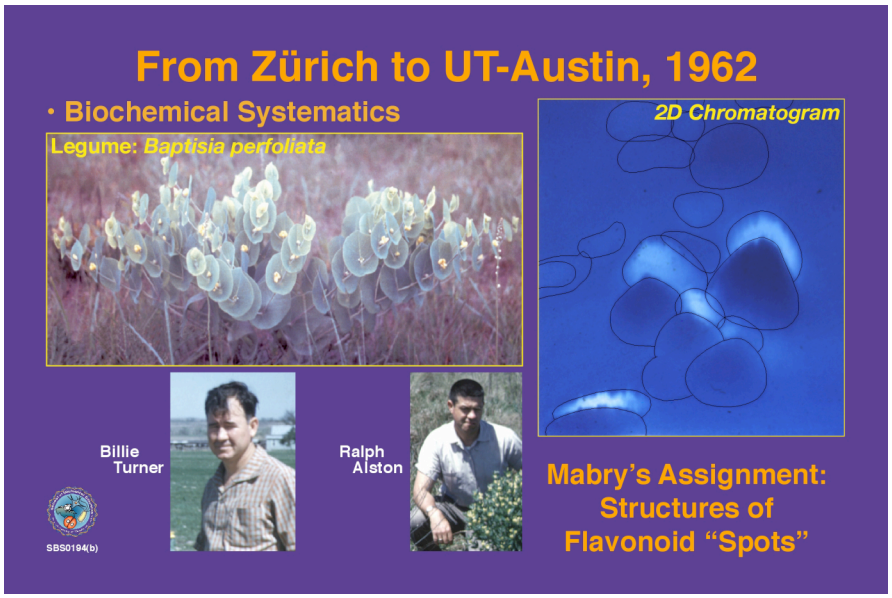


Figure 3. The biochemical systematic investigations underway at UT-Austin in 1962 utilized 2D paper chromatography to distinguish species and recognize hybrids in the legume genus *Baptisia* as well as other taxa. Mabry's phytochemistry laboratory was charged with providing structures for the florescent flavonoid "spots" detected when the chromatograms were viewed under UV light.

Figure 4. Dr. Mabry (center) with coworkers in the late 1960's, including his future wife Helga Fischer (center right), a Chemical Technician from Germany. Coworkers, clockwise from top left: Dr. Ken Markham, Al Wohlpart*, Gene Miller*, Julis "Bud" Kroschewsky*, Helga, Hanspeter Rüesch (Chem. Tech), Dr. Klaus Fischer, Dr. Michael Thomas, Chistina Chang*; * Ph.D. students

By 1970, Dr. Mabry and coworkers (see Figure 4) had documented the flavonoid patterns from an enormous number of plants from around the world. In addition to publishing many papers on their findings, they produced what quickly became the standard text for flavonoid studies, the book titled "The Systematic Identification of Flavonoids" [8]. This volume, which is still used today, includes 1000's of interpreted UV and NMR spectra of members of different classes of flavonoid aglycones and glycosides. Moreover, the book presents details of how to best extract and purify flavonoids from plants in a user-friendly format that allows botanists with little or no chemistry training to identify flavonoids and other phenolic compounds using spectral techniques (Figure 5).

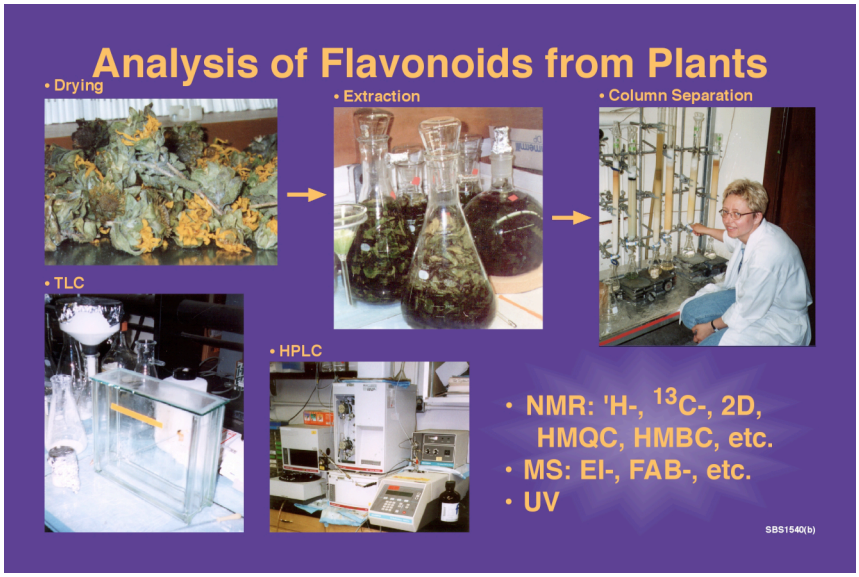


Figure 5. Some of the analysis procedures developed for flavonoids in Mabry's lab

Today, plant systematics is dominated by DNA studies; nevertheless, some plant relationships are still better understood by combining chemical and molecular data; for example, in the 1960-1970's, with the help of the plant systematists at UT-Austin, Mabry and coworkers established that betalains were restricted to some nine plant families in the Order Caryophyllales, some aligned solidly there for the first time (e.g. the Cactaceae). Mabry's group later used DNA studies to confirm the close evolutionary relationship of all these families; however, these nine betalain families contained no anthocyanins (pigments that are in found over 95% of flowering plants) and somewhat to Mabry's surprise, two other families that clearly belong in the Order Caryophyllales, the Caryophylaceae and Molugenaceae, contained only anthocyanins and no betalains (Figure 6) [9, 10, 11, 12, 13, 14, 15, 16 and many other Mabry papers].

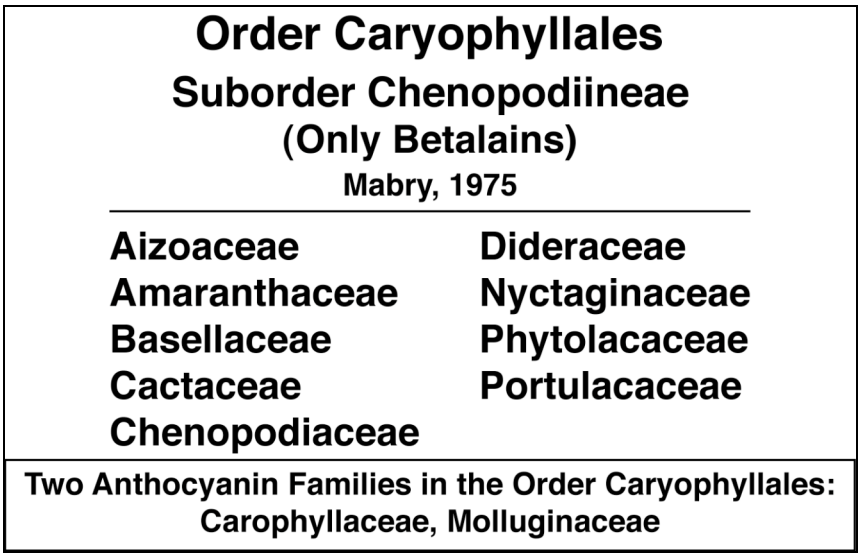


Figure 6. In higher plants, betalains were found to occur in nine plant families in the Order Caryophyllales, which also houses two families that have only anthocyanins, that is, no betalains.

This example illustrates both the usefulness and the limitations of chemical data in understanding plant evolutionary relationships. Mabry's group found that in addition to systematics, natural products chemistry can often aid in resolving many plant biology questions, such as determining the survival potential of plants in particular ecosystems [17, 18, 19 and many other Mabry papers].

Also in the 1960's-1970's, Mabry expanded his plant chemistry program to cover many other classes of natural products, for example, different types of phenolic compounds as well as sesquiterpene lactones and diterpenes, studies that resulted in many papers as well as two data-filled volumes [20, 21]. Other groups of natural products investigated include all other classes of terpenes, as well as nonprotein amino acids, azoxyglycosides, some alkaloids and so on (many Mabry papers). Mabry also investigated the chemistry of numerous disjunct taxa, including populations of some species found in North and South America [22, 23, 24, 25] as well as the chemistry of many arid land plants [26, 27, 28, 29, 30, 31]. In addition, his group tested many of the new natural products for a wide range of biological properties including their antifeedant, antibiotic, anticancer, free radical scavenging, anti-inflammatory, hepatoprotective and

other activities to determine some of their roles in evolution as well as their potential for medical uses.

Dr. Mabry always encouraged students who presented a good research plan to tackle projects of their own interest, even when outside his areas of specialization. For example, Ph.D. student Maureen Bonness, who, in the 1980's, at Dr. Mabry's suggestion initiated investigations of betalains in pokeweed (*Phytolacca americana*, family Phytolacaceae). However, while reading literature on this plant, Bonness learned that this very same species contains a special class of proteins known as "PAP" for "pokeweed antiviral protein" (the more general name is "RIP" for "ribosome inactivating protein"). So Bonness decided to investigate the antiviral proteins in *Phytolacca americana* and quickly confirmed that PAP can inactivate ribosomes inside the *Phytolacca* cells, preventing any "PAP-invaded" cells from replicating [32]. Bonness with other coworkers [e.g. 33] established that PAP is in an active form in the cell walls of this *Phytolacca* species, and, in addition, she found that when a virus enters through the cell wall, PAP also enters and that the PAP quickly inactivates the ribosomes, which not only causes the death of these PAP-invaded cells but also prevents replication of the virus, keeping the remainder of the plant free of the virus (Figure 7) [32]. Other laboratories have reported that when the gene for PAP is inserted into another plant, such as those for crops, these genetically modified plants are protected from viruses; unfortunately, the next generations of these genetically modified plants usually eliminate the gene for PAP. Also methods are currently being developed in European labs and elsewhere to use PAP to protect human cells against the HIV virus and some cancer viruses. Mabry hopes these remarkable antiviral proteins will continue to be investigated for both agricultural and medical uses.

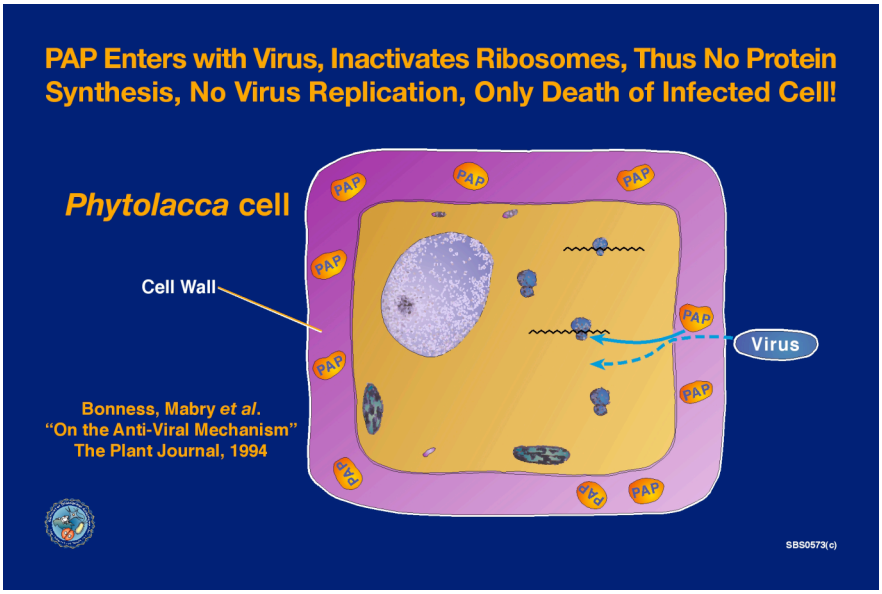


Figure 7. The mechanism by which the antiviral proteins (PAP) that occur in the cell walls of *Phytolacca americana* protect these plants from viral infections; PAP "piggybacks" into a *Phytolacca* cell when a virus enters and then kills that single infected cell, thus preventing spread of the virus.

Bonness's curiosity and keen observation skills also helped to establish another major project in the Mabry lab. When she happened to notice in a class project that fungus-contaminated cell cultures of the cactus *Cephalocerus senilis* (old man cactus) turned red, she quickly established that the red pigments were not betalains. This finding lead to further investigations by others in Mabry's group, namely Paul Paré and coworkers, who found that the new compounds in these cultures were being elicited by chitin (an elicitor of protective compounds for plants that was known to occur in fungi). Paré and others in the Mabry lab established that most of the elicited compounds belonged to a new class of 4′-deoxyflavonoids, including an unusual new aurone with high antibiotic properties (Figure 8) [34, 35, 36, 37, 38, 39]. Cultures of another distantly related cactus, *Opuntia microdasys*, also turned red when treated with chitin and preliminary analysis showed a similar pattern of 4′-deoxyflavonoids [40]. Mabry believes that treating cultures of other cacti, as well as plants from other families, with elicitors such as chitin could produce many more new and important compounds.

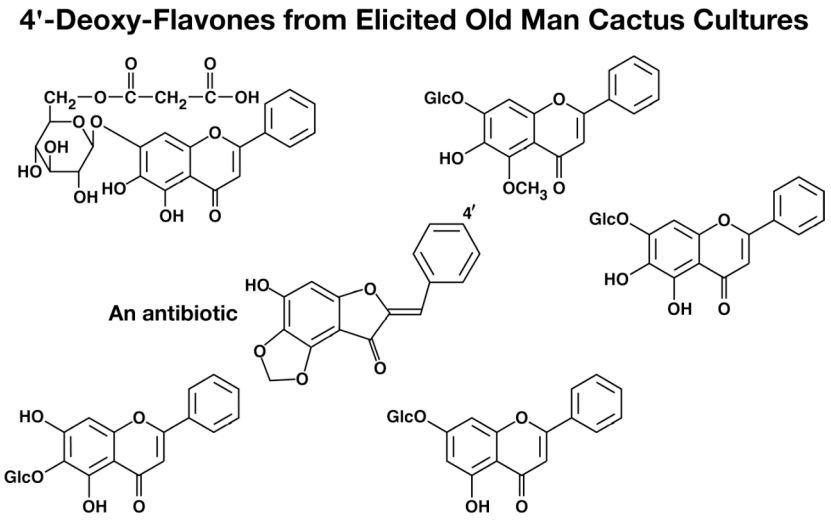


Figure 8. Fungus-infected cactus cell cultures produced unique 4′-deoxyflavonoids as a result of the action of the elicitor chitin, a fungal component.

In the late 1980's Dr. Len Kurland, a distinguished ethnobotanist who spent his early years at NIH before joining the Mayo Clinic in Rochester, Minnesota, presented a new project to Mabry. Kurland pointed out to Mabry that a terrible neurodisease occurring in members of the Chamorro tribe on the island of Guam exhibited symptoms of ALS, Parkinson's dementia and Alzheimer's and was over 100 times more prevalent on Guam than such neurodiseases elsewhere. After years of studying this disease and finding no other cause for it, Kurland finally focused on a possible dietary "culprit." Kurland explained to Mabry that the Chamorros traditionally consumed seeds of the Guam cycad (*Cycas micronesia*) leading him and others to suspect that these seeds might contain a neurotoxin. But other scientists determined that there were only very small quantities of the "suspect" compound, β-methylaminoalanine (BMAA), in the thoroughly washed cycad flour being used for tortillas by the Chamorros. Kurland persuaded Mabry to reinvestigate whether cycads and BMAA (or other compounds) were involved in this Guam neurodisease. Delia Brownson, an excellent Ph.D. student, initiated the research by searching literature reports on the project, and learned that neurodiseases were, in many cases, caused by channels in human brain cells opening improperly, allowing too much calcium into the cells, which then caused the death of some of the cells, leading to neurodiseases. She also learned that BMAA, the Guam cycad compound, complexes

with CO_2 in solutions of BMAA and CO_2 at physiological pH's to form an unstable compound called a carbamate. This unstable carbamate (present in about 10% yield in an equilibrium reaction in these solutions) has a structure remarkably similar to the neurotransmitter glutamic acid, one of the very compounds that our brains use to properly and regularly manipulate calcium channels. Suspecting involvement of the carbamate of BMAA, Brownson first tested BMAA alone in solutions of fresh rat brain cells and found that, as expected, no calcium entered the cells. Next, she added sodium bicarbonate (a source of CO_2) to the same solution of BMAA and rat brain cells, and, sure enough, the channels opened and calcium entered the rat brain cells (Figures 9, 10) [41,42]. These results suggest that the carbamate of BMAA may be the cause of the neurodisease that afflicted so many Chamorros on Guam.

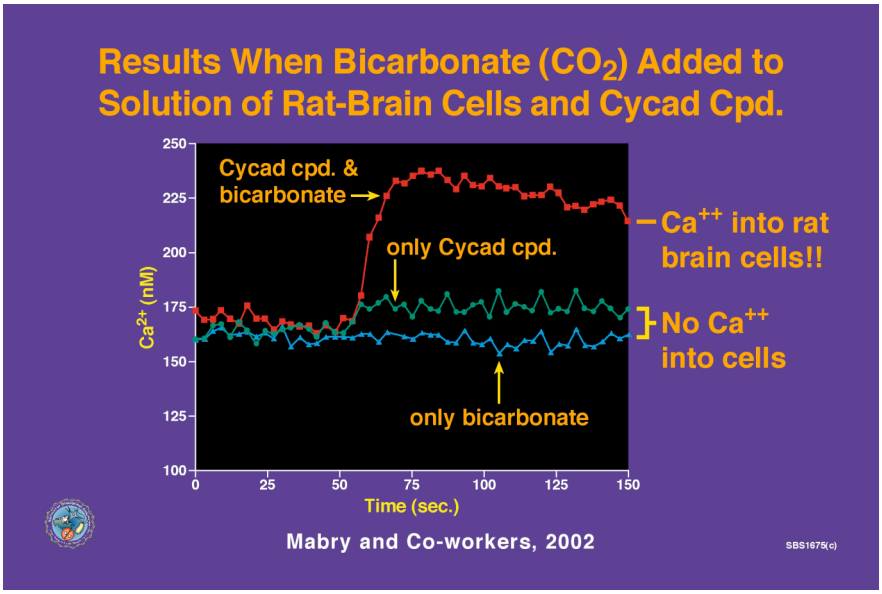


Figure 9. Mabry's lab showed that the cycad compound β-methylaminoalanine (BMAA) could increase Ca^{2+} levels in rat brain cells, but only in the presence of bicarbonate, implicating the carbamate of BMAA as the cause of ALS-PDC.

Increasing the significance of Brownson's findings were reports in this decade by Paul Cox and his coworkers (earlier based at the Institute of Ethnobotany in Kalaheo, Hawaii, but now associated with Brigham Young University, Provo, Utah) who proposed what Mabry believes is the probable route by which BMAA enters the brains of the

Chamorros: Cox's group established that Guam bats accumulate BMAA in their bodies by eating the cycad seeds and that the Chamorros feast on these BMAA-rich bats [43 and later publications]!

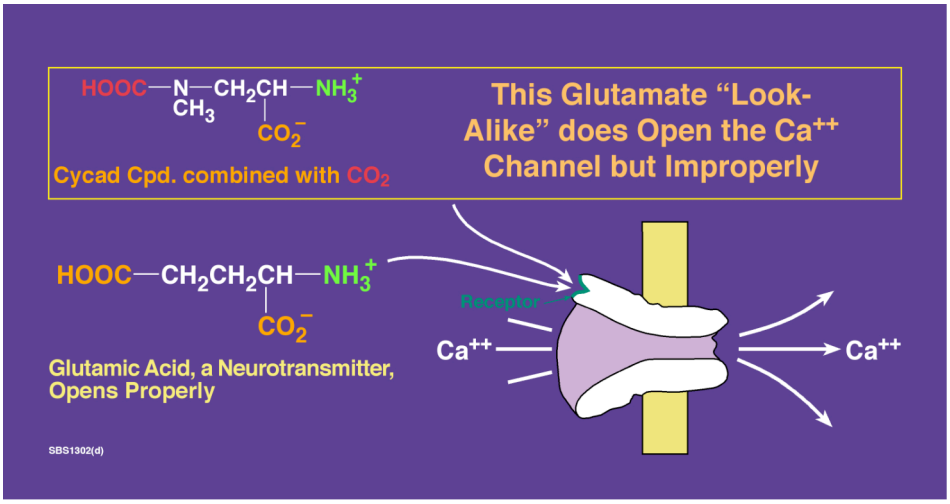


Figure 10. BMAA, the cycad compound, can complex with CO_2 to give a carbamate. Mabry's group showed that this carbamate allows calcium to enter rat brain cells, presumably by binding to the glutamate receptor protein and opening the calcium channels. Mabry believes this carbamate may be the cause of ALS-PDC.

Although this theory has received some skepticism, there is other evidence that points to the dietary consumption of bats as the likely route of BMAA into the brains of the Chamorros, namely, when in the early 1970's the Guam bats were classified as an endangered species and hunting of them was outlawed, no new cases of the neurodisease were reported on Guam!! For obvious reasons, Dr. Mabry considers that it is extremely urgent for researchers to finally determine if compounds such as the carbamate of BMAA can indeed cause neurodiseases by binding to receptor proteins for neurotransmitters and improperly opening the calcium channels.

Now I mention my own Ph.D. studies that are being supervised by Dr. Mabry. My dissertation research involves both phytochemical and chemosystematic studies of members of one of the largest and most diverse plant families in the world, the Sunflowers. The genus *Silphium* L. (Asteraceae) consists of eleven species native to North America, distributed primarily in the eastern United States and extending into

southeastern Canada [44]. I became interested in this genus because extracts of several *Silphium* species were widely used for herbal medicines by central and southeastern Native American tribes [45], yet there have been few attempts to correlate phytochemical findings with their traditional herbal uses. One previous study led to the isolation and characterization of eight major triterpene glycosides from *Silphium perfoliatium* (Figure 11) [46]. Also, a mixture of these triterpene saponins from *S. perfoliatum* showed blood cholesterol-lowering activity, as well as fungicidal properties [47, 48].

Thus, in an effort to further characterize triterpene glycosides from *Silphium* and determine some of their medical importance, I initiated an analysis of saponins in all species of *Silphium*, some of which contain up to nine triterpene saponins with new algycones.

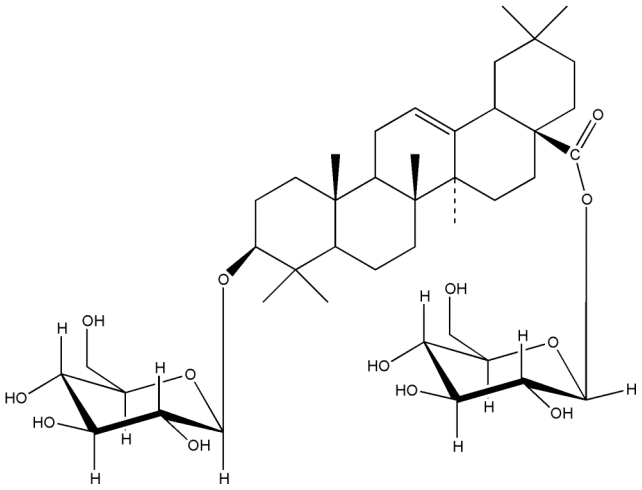


Figure 11. Silphioside B, one of the 8 major triterpene saponins previously isolated from *Silphium perfoliatum* [46].

Since the investigation of saponins was an entirely new area of research in the Mabry lab, the first step in my dissertation project was to gain instrumentation training and learn isolation techniques to characterize saponins. With Dr. Mabry's support and encouragement, I established connections with a number of leaders in the field of saponin chemistry, particularly, in the laboratory of Professor Wieslaw Oleszek in Pulawy, Poland where I learned procedures essential for the characterization of saponins. In

addition, I gained skills for interpreting complex NMR spectra of my compounds from Dr. Ben Shoulders, Professor of Chemistry at UT-Austin and from Professor Sonia Piacente, a NMR specialist for saponins at the University of Salerno, Italy. Once the major saponins from *Silphium* were characterized, the cytotoxic activities of the isolated compounds were tested against human breast cancer cell lines in collaboration with Dr. Su Dharmawardhane, Universidad Central del Caribe School of Medicine, Puerto Rico. The results indicate that one of the new saponins isolated from *Silphium radula* decreased proliferation of breast cancer cells in a statistically significant manner. Finally, I was able to establish an LC-MS detection method for saponins in *Silphium* in order to identify the major saponins in all *Silphium* species with the help of Dr. Paul Paré and coworkers at Texas Tech University, Lubbock, Texas. My analyses of the many new and known triterpene saponins in all eleven species of *Silphium*, as well as my findings for their systematic and medicinal value, will be complete and published in the 2007-2008 academic year.

During Dr. Mabry's retirement celebration in October 2005 at UT-Austin, there were three major speakers, all former students of Dr. Mabry: Dr. Paul Paré and Dr. Jonathan Gershenzon, as well as Dr. Barbara Timmermann, Chair of the Department of Medicinal Chemistry and Distinguished Professor of Medicinal Chemistry, School of Pharmacy, University of Kansas. Dr. Timmermann mentioned in her lecture that once when Dr. Mabry had helped her resolve a difficult issue, she asked what she could do for him, he responded "pass it on," meaning that in her role as a professor, she must not only educate and train students with all her knowledge and techniques for this field but also help them develop essential people-to-people skills. Moreover, she must inspire and support students through their challenging times so they can achieve their goals and dreams. Later in the celebration, Dr. Jim Gill (M.D.) pointed out that he might not have reached his goal of becoming a medical doctor without the opportunity to spend his undergraduate years in Mabry's lab where he learned not only biological chemistry but also the skills to operate Mabry's MS and NMR equipment. Next, Dr. Gill announced that an endowment had been created ("Professor Tom J. Mabry Endowed Excellence

Fund in Phytochemistry and Plant Biology") to provide awards for research excellence by outstanding graduate students in the plant sciences.

In a letter to me, Dr. Bonness who finished in 1992, described Dr. Mabry as "undeniably the BEST major professor that anyone could have wanted, always supporting his students above and beyond the call of duty". She recalled how he championed his students for awards, TA positions, fellowships and so many other opportunities. And when he couldn't provide advice himself he arranged meetings with key people in that particular area of research, which often led to successful collaborations. Now, almost twenty years later, my personal experience as Dr. Mabry's last graduate student echoes Dr. Bonness's feelings in every way. Of course, the lab is a much different place now since it is nearing closure, but I still feel lucky to be able to enjoy everyday his untamed enthusiasm for science, for teaching and for life in general (Figure 12).

Figure 12. Dr. Mabry and his last Ph.D. student Lalita Calabria (senior author of this paper) in Lublin, Poland, June 2006, at the 5th International Symposium on Chromatography of Natural Products.

From the many individuals who have known and had the pleasure of working closely with Dr. Mabry, including myself, we would like to express our deep appreciation and affection for him, especially on the occasion of his 75th birthday. There is no question that his career-long dedication to outstanding research and teaching excellence has impacted the field of phytochemistry and inspired many lives forever.

References

[1] Mabry TJ. (**1960**) Syntheses and Coenzymatic Activities of Analogs of Vitamin C. *Dissertation*. *Rice University, Houston*.

[2] Ettlinger MG, Dateo GP, Harrison BW, Mabry TJ, Thompson CP. (**1961**) Vitamin C as a coenzyme: the hydrolysis of mustard oil glucosides. *Proceedings of the National Academy of Sciences*, ***47***, 1875-1880.

[3] Mabry TJ, Wyler H, Sassu G, Mercier M, Parikh I, Dreiding AS. (**1962**) Die Struktur des neobetanidins. *Helvetica Chimica Acta, **45**,* 640-647.

[4] Wyler H, Mabry TJ, Dreiding AS. (**1963**) Zur Struktur des Betanidins: über die Konstitution des Randenfarbstoffes Betanin. *Helvetica Chimica Acta, **46***, 1745-1748.

[5] Mabry TJ, Dreiding AS. (**1968**) The betalains. In *Recent Advances in Phytochemistry, Vol. 1*. TJ Mabry, RE Alston and V.C. Runeckles (Eds), Appleton-Century-Crofts, New York, 145-160.

[6] Alston R, Turner BL. (**1963)** *Biochemical Systematics.* Prentice Hall, Inc., New Jersey, 404.

[7] Mabry TJ, Alston RE, Turner BL. (**1967**) Chemotaxonomy. In *Encyclopedia of Biochemistry.* R.J. Williams and E.M. Lansford, Jr. (Eds), New York, 226-228.

[8] Mabry TJ, Markham KR, Thomas MB. (**1970**) *The Systematic Identification of Flavonoids*. Springer-Verlag, Heidelberg-New York, 354.

[9] Mabry TJ, Taylor A, Turner BL. (**1963**) The betacyanins and their distribution. *Phytochemistry*, **2**, 61-64.

[10] Mabry TJ, Turner BL. (**1964**) Chemical investigations of the Batidaceae, betaxanthins and their systematic implications. *Taxon, **13***, 197-200.

[11] Mabry TJ. (**1966**) The betacyanins and betaxanthins. In *Comparative Phytochemistry*. T. Swain (Ed), Academic Press, London, 231-244.

[12] Mabry TJ. (**1970**) The betalains. In *The Chemistry of Alkaloids*. S.W. Pelletier (Ed), Reinhold, 367-384.

[13] Mabry TJ. (**1973**) Is the Centrospermae monophyletic? A review of phylogenetically significant molecular, ultrastructural and other data for Centrospermae families. In *Chemistry in Botanical Classification, Nobel Symposium No. 25*. G. Bendz, J. Santesson and V. Runnstrom-Reio (Eds), Nobel Foundation, Stockholm, Sweden and Academic Press, New York, 75-85.

[14] Wohlpart A, Mabry TJ. (**1968**) The distribution and phylogenetic significance of the betalains with respect to the Centrospermae. *Taxon, **17,*** 148-152.

[15] Behnke HD, Mabry TJ. (Eds). (**1994**) *Caryophyllales: Evolution and Systematics*. Springer-Verlag, Heidelberg, 276.

[16] Mabry TJ, Behnke HD. (Eds). (**1976**) *The Evolution of Centrospermous Families.* Springer-Verlag, Vienna, 110.

[17] Gershenzon J, Mabry TJ. (**1983**) Secondary metabolites and the higher classification of angiosperms. *Nordic Journal of Botany, **3***, 5-34.

[18] Rossiter MC, Gershenzon J, Mabry TJ. (**1986**) Behavioral and growth responses of the specialist herbivore, *Homoeosoma electellum*, to major terpenoid of its host, *Helianthus* spp. *Journal of Chemical Ecology, **12,*** 1505-1521.

[19] Rogers CE, Gershenzon J, Ohno N, Mabry TJ, Stipanovic RD, Kreitner GL. (**1987**) Terpenes of wild sunflower (*Helianthus*): an effective mechanism against seed predation by larvae of the sunflower moth, *Homoeosoma electellum* (Lepidoptera: Pyralidae). *Environmental Entomology*, ***16,*** 586-592.

[20] Yoshioka HT, Mabry TJ, Timmermann B. (**1973**) *Sesquiterpene Lactones*. The University of Tokyo Press, Tokyo, 544.

[21] Seaman F, Bohlmann F, Zdero C, Mabry TJ. (**1989**) *Diterpenes of Flowering Plants, Compositae (Asteraceae)*. Springer-Verlag, New York, 638.

[22] Potter JL, Mabry TJ. (**1972**) Origin of the Texas Gulf coast island populations of *Ambrosia psilostachya* DC. (Compositae): A numerical study using terpenoid data. *Phytochemistry, **11***, 715-723.

[23] Mabry TJ. (**1973**) The chemistry of geographical races. In *Chemistry in Evolution and Systematics*. T. Swain (Ed), Butterworth, London, 377-400.

[24] Mabry TJ. (**1973**) The chemistry of disjunct taxa. In *Chemistry in Botanical Classification, Novel Symposium No. 25*. G. Bendz, J. Santesson and V. Runnstrom-Reio (Eds), Nobel Foundation, Stockholm, Sweden, and Academic Press, New York, 63-66.

[25] Carman NJ, Mabry TJ. (**1975**) Disjunction of *Prosopis reptans*, and the origin of the North American populations. *Biochemistry Systematics and Ecology, **3***, 19-23.

[26] Mabry TJ, Hunziker J, Difeo DR. (Eds). (**1977**) *The Creosote Bush: The Biology and Chemistry of Larrea in the New World Deserts*. Dowden, Hutchinson and Ross, Inc., Stroudsburg, Pennsylvania, 284.

[27] Campos-Lopez E, Mabry TJ, Fernandez S. (Eds). (**1979**) *Larrea, a Vast Resource of the American Deserts*. Centro de Investigacion en Quimica Aplicada, Saltillo, Coahuilla, Mexico, 411.

[28] Mabry TJ, Difeo DR. (**1973**) The role of the secondary plant chemistry in the evolution of the Mediterranean scrub vegetation. In *Ecological Studies. Analysis and Synthesis, Vol. 7*. F. di Castri and H.A. Mooney (Eds), Springer-Verlag, Heidelberg-New York, 121-155.

[29] Mabry TJ, DiFeo DR, Sakakibara M, Bohnstedt CF, Seigler D. (**1977**) The natural products chemistry of *Larrea*. In *The Creosote Bush: The Biology and Chemistry of Larrea in the New World Deserts*. T.J. Mabry, J. Hunziker and D.R. DiFeo, Jr. (Eds), Dowden, Hutchinson and Ross, Inc., Stroudsburg, Pennsylvania, 115-134.

[30] Mabry TJ, Bohnstedt CF, Jr. (**1979**) Larrea: a chemical resource. In *Larrea: A Vast Resource of the American Deserts*. E. Campos-Lopez, T.J. Mabry and S. Fernandez (Eds), Centro de Investigacion en Quimca Aplicada, Saltillo, Coahuilla, Mexico, 217-236.

[31] Mabry TJ, Nguyen H, Dixon RA, Bonness MS. (Eds). (**1993**) *The Biotechnology of Aridland Plants*. IC2 Institute, UT-Austin, 369.

[32] Bonness MS, Ready MP, Irvin JD, Mabry TJ. (**1994**) Pokeweed antiviral protein inactivates pokeweed ribosomes; implications for the antiviral mechanism. *The Plant Journal, **5**,* 173-183.

[33] Lu H. (**1996**) Ribosome-Inactivating Proteins from *Phytolacca* Leaves are Sequestered in the Cell Wall in an Active Form. *Master's Thesis. The University of Texas at Austin.*

[34] Paré PW, Mischke CF, Edwards R, Dixon RA, Norman HA, Mabry TJ. (**1992**) Induction of phenylpropanoid pathway enzymes in elicitor-treated cultures of *Cephalocereus senilis*. *Phytochemistry*, ***31***, 149-153.

[35] Bonness MS, Paré PW, Mabry TJ. (**1993**) Novel callus and suspension cultures of the "old man" cactus (*Cephalocereus senilis*). *Cactus and Succulent Journal, **65**,* 144-147.

[36] Liu Q, Markham KR, Paré PW, Dixon RA, Mabry TJ. (**1993**) Flavonoids from elicitor-treated cell suspension cultures of *Cephalocereus senilis*. *Phytochemistry, **32**,* 925-928.

[37] Liu DQ. (**1994**) Chemical and Enzymological Investigations of the Phenylpropanoid Pathway in Elicited Cultures of *Cephalocereus senilis* (old man cactus). *Dissertation. The University of Texas at Austin.*

[38] Padolina I. (**2002**) Chitin-Induced Biosynthesis of Phytoalexin 4'-Deoxyaurone in Cell Suspension Cultures of "Old Man" Cactus, *Cephalocereus senilis*. *Dissertation. The University of Texas at Austin.*

[39] Han M. (**2004**) Identification of the Compounds from Chitin-Elicited Cultures of *Cephalocereus senilis* by LC/MS/MS (Liquid Chromatography/Ion Trap Mass Spectrometry/Triple Quadruple Mass Spectrometry). *Master's Thesis. The University of Texas at Austin.*

[40] Zou Y. (**2006**) Chemical Comparison of the Chitin-Induced (Elicited) Flavonoid Pathway in Cell Cultures of Old Man Cactus and in Bunny Ears Cactus by ESI-MS. *Master's Thesis. The University of Texas at Austin.*

[41] Brownson DM. (**1996**) A Study of Known Excitotoxic Compounds and Isolated Nonprotein Amino Acids from Cycads. *Dissertation. The University of Texas at Austin.*

[42] Brownson DM, Mabry TJ, Leslie SW. (**2002**) The cycad neurotoxic amino acid, ß-*N*-methylamino-L-alanine (BMAA), elevates intracellular calcium levels in dissociated rat brain cells. *Journal of Ethnopharmacology, **82**,* 159-167.

[43] Cox PA, Sacks OW. (**2002**) Cycad neurotoxins, consumption of flying foxes, and ALS-PDC disease in Guam: a medical hypothesis. *Neurology, **58**,* 956-959.

[44] Clevinger, JA, Panero, JL. **(2000)** Phylogenetic analysis of *Silphium* and subtribe Engelmanniinae (Asteraceae:Heliantheae) based on ITS and ETS sequence data. *American Journal of Botany,* **87**, 565-572.

[45] Hamel, PH, Chiltoskey, MU. **(1975)** Cherokee plants and their uses: A four hundred year history. Herald Publishing Co., North Carolina

[46] Davidyants, ES, Putieva, ZhM, Bandyakova VA, Abubakirov, NK. **(1984)** Triterpene glyocosides from *Silphium perfoliatum. Khimiya Prirodnykh Soedinenii*, **1**, 120-121.

[47] Syrov, VN, Khushbaktova, ZA, Davidyants, ES. **(1992)** Triterpenoid glycosides from *Silphium perfoliatum* L.: Hypolipidemic activity of Silphioside A. *Khim-Farm. Zh.* ***26***, 66-69.

[48] Davidyants, ES, Kartasheva, IA, Neshin, IV. **(1997)** The effect of triterpene glycosides from *Silphium perfoliatum* L. on phytopathogenic fungi. *Rastitel'ne Resursy*, **33**, 93-98.

References

Achine, L., Huhman, D.V., Farag, M.A., Sumner, L.W., Blount, J.W., Dixon, R.A., 2005. Genomics-based selection and functional characterzation of triterpene glycosyltranferases from the model legume *Medicago truncatula*. Plant J. 41, 875-887.

Akihisa, T., Yasukaw, K., Hirotoshi, O., Yoshimasa, K., Yamanouchi, S., Takido, M., Kumaki, K., Tamura, T., 1996. Triterpene alcohols from the flowers of Compositae and their anti-inflammatory effects. Phytochemistry 43, 1255-1260.

Alvarenga, S.A.V., Ferreira, M.J.P., Emerenciano, V.P. & Cabrol-Bass, D., 2001. Chemosystematic studies of natural compounds isolated from Asteraceae. Characterization of tribes by principal component analysis. Chemom. Intell. Lab. Syst. 56, 27-37.

Alvarenga, S.A.V., Ferreira, M.J.P., Rodrigues, G.V., Emerenciano, V.P., 2005. A general survey and taxonomic implications of diterpenes in the Asteraceae. Bot. J. Linn. Soc. 147, 291-308.

Amimoto, K., Yoshikawa, K., Arihara, S., 1993. Triterpenoid saponins of aquifoliaceous plants. XII. Ilexosides XLVI-LI from the leaves of *Ilex rotunda* Thunb. Chem. Pharm. Bull. 41, 77-80.

Anke, S., Niemüller, D., Moll, S., Hänsch, R., Ober, D., 2004. Polyphyletic origin of pyrrolizidine alkaloids within the Asteraceae. Evidence for differential tissue expression of homospermidine synthase. Plant Physiol. 136, 4437-4047.

Araujo, D.S., Chaves, M.H., 2005. Pentacyclic triterpenoids from leaves of *Terminalia brasiliensis*. Quim. Nova 28, 996-999.

Asfaw, N., Storesund, H.J., Skattebol, L., Aasen, A.J., 1999. (1S,5R)-(-)-2,4,4-trimethylbicyclo[3.1.1]hept-2-en-6-one, from the essential oil of the Ethiopian plant *Laggera tomentosa*. Phytochemistry 52, 1491-1494.

Asongalem, E.A., Foyet, H.S., Ngogang, J., Folefoc, G.N., Dimo, T., Kamtchouing, P., 2004. Analgesic and anti-inflammatory activities of *Erigeron floribundus*. J. Ethnopharmacol. 91, 301-308.

Bader, G., Plohmann, B., Hiller, K., Franz, G., 1996. Cytotoxicity of triterpenoid saponins. Part 1: Activities against tumor cells in vitro and hemolytical index. Pharmazie 51, 414-417.

Basyuni, M., Oku, H., Inafuku, M., Baba, S., Iwasaki, H., Oshiro, K., Okabe, T., Shibuya, M., Ebizua, Y., 2006. Molecular cloning and functional expression of a multifunctional triterpene synthase cDNA from a mangrove species *Kandelia candel* (L.) Druce. Phytochemistry 67, 2517-2524.

Bayer, R.J, Starr, J.R., 1998. Tribal phylogeny of the Asteraceae based on two non-coding chloroplast sequences, the trnL intron and trnL/trnF intergenic spacer. Ann. Mo. Bot. Gard. 85, 242-256.

Bentham, G., 1873. Notes on the classification, history, and geographical distribution of the Compositae. Bot. J. Linn. Soc. 13, 335-557.

Berendsohn, W.G., Jakupovic, J., Zdero, C., 1998. Natural Substances in the Compositae: The Bohlmann Files. [http://www.bgbm.org/BioDivInf/projects/bohlmannfiles/]

Blay, G., Garcia, B., Molina, E., Pedro, J.R., 2006. Syntheses of (+)-alismoxide and (+)-4-epi-alismoxide. J. Org. Chem. 71, 7866-7869.

Bohlmann F., Jakupovic, J., 1979. Naturally occurring terpene derivatives. 232. New labdane derivative and sesquiterpene in *Silphium* species. Phytochemistry 18, 1987-1992.

Bohlmann, F., Jakupovic, J., 1980. Natural terpene derivatives. 238. New sesquiterpene hydrocarbons with anomalous carbon skeletons from *Silphium* species. Phytochemistry 19, 259-265.

Bohm, B.A., 1998. *Introduction to flavonoids.* Harwood Academic Publishers, Amsterdam.

Bohm, B.A., Reid, A., DeVore, M., Stuessy, T.F., 1995. Flavonoid chemistry of the Calyceraceae. Can. J. Bot. 73, 1962-1965.

Bohm, B.A., Stuessy, T.F., 2001. *Flavonoids of the sunflower family (Asteraceae).* Springer-Verlag/Wien, New York.

Botanical Society of America,1995. Botany for the next millennium; a report from the Botanical Society of America. Columbus, Ohio, USA.

Bremer, K., 1987. Tribal interrelationships of the Asteraceae. Cladistics 3, 210-253.

Bremer, K., 1994. Asteraceae - Cladistic & Classification, Timber Press, Portland, OR, 1-752.

Bremer, K., 1996. Major clades and grades of the Asteraceae. In Compositae: Systematics In: Compositae: Systematics. Proceedings of the International Compositae Conference, Kew, 1994. Hind DJN (Editor-in-Chief) Vol. 1. Royal Botanic Gardens, Kew, 1-7.

Calabria, L.M., Piacente, S., Kapusta, I., Dharmawardhane, S.F., Segarra, F.M., Pessiki, P.J., Mabry, T.J., 2008. Triterpene saponins from *Silphium radula*. Phytochemistry 69, 961-972.

Calabria L.M., Emerenciano, V.P., Scotti, M.T., Mabry, T.J., 2007. Phylogenetic analysis of tribes of the Asteraceae based on phytochemical data. Nat. Prod. Comm. 3, 277-285.

Calabria, L.M., Emerenciano, V.P., Scotti, M.T., Mabry, T.J., 2006. Secondary Chemistry of the Compositae (Asteraceae). In The International Compositae Alliance Meeting Book of Abstracts. Barelona, Spain.

Carlquist, S., 1976. Tribal interrelationships and phylogeny of the Asteraceae. Aliso 8, 465-492.

Cassini, H., 1816. Tableau exprimant les affinities des tribus naturelles de famile des Synanthérées. In Dictionnaire des Sciences Naturelles, vol. 3. Ed. G. Cuvier. 2nd ed. Paris: Le Normant.

Chaudhuri, P.K., 1992. 7-hydroxyechinozolinone, a new alkaloid from the flowers of *Echinops echinatus*. J. Nat. Prod. 55, 249-250.

Chen, J.C., Chang, N.W., Chung, J.G., Chen, K.C., 2003. Saikosaponin-A induces apoptotic mechanism in human breast MDA-MB-231 and MCF-7 cancer cells. Am. J. Chin. Med. 31, 363-377.

Christensen, L.P., Lam, J., 1990. Acetylenes and related-compounds in Cynareae. Phytochemistry 29, 2753-2785.

Christensen, L.P., Lam, J., 1991a. Acetylenes and related-compounds in Heliantheae. Phytochemistry 30, 11-49.

Christensen, L.P., Lam, J., 1991b. Acetylenes and related-compounds in Asteraceae (= Compositae)3. Acetylenes and related-compounds in Astereae. Phytochemistry 30, 2453-2476.

Christensen, L.P., 1992. Acetylenes and related-compounds in Asteraceae (= Compositae) .4. Acetylenes and related-compounds in Anthemideae. Phytochemistry 31, 7-49.

Clevinger, J.A., 1999. Systematics of *Silphium* and its subtribe Engelmanniinae (Asteraceae: Heliantheae). Texas, USA: University of Texas at Austin (PhD Dissertation).

Clevinger, J.A., 2004. Four new combinations in *Silphium*. Novon 14, 275-277.

Clevinger, J.A., 2006. *Silphium*. Flora of North America, North of Mexico. Vol. 21 Asteraceae. Flora of North America Editorial Committee, Oxford University Press, New York and Oxford. pp. 77-82.

Clevinger, J.A., Panero, J.L., 2000. Phylogenetic analysis of *Silphium* and subtribe Engelmanniinae (Asteraceae: Heliantheae) based on ITS and ETS sequence data. Am. J. Bot. 87, 565-572.

Cook, W., 1869. The Physiomedical Dispensatory. Henriette's Herbal Homepage: [http://www.henriettesherbal.com/]

Corea, G., Iorizzi, M., Lanzotti, V., Cammareri, M., Conicella, C., Laezza, C., Bifulco, M., 2004. Astersedifolioside A-C, three new oleanane-type saponins with antiproliferative activity. Bioorg. Med. Chem. 12, 4909-4915.

Cronquist, A., 1980. Vascular Flora of the Southeastern United States. The University of North Carolina Press. 1, 67-69.

D'Auria, J.C., Gershenzon, J. 2005. The secondary metabolism of *Arabidopsis thaliana*: growing like a weed. Curr. Opin. Plant Bio. 8, 308-316.

Davidyants, E.S. 2006. Growth-regulating activity of triterpene glycosides from *Silphium perfoliatum* (Asteraceae). Rastit. Resur. 42, 127-136.

Davidyants, E.S., Kartasheva, I.A., Neshin, I.V., 1997. The effect of triterpene glycosides of *Silphium perfoliatum* L. on phytopathogenic fungi. Rastit. Resur. 33, 93-98.

Davidyants, E.S., Putieva, Zh.M., Bandyukova, V.A., Abubakirov, N.K., 1984a. Triterpene glycosides of *Silphium perfoliatum*. Chem. Nat. Compd. 1, 120-121.

Davidyants, E.S., Putieva, Zh.M., Bandyukova, V.A., Abubakirov, N.K., 1984b. Triterpene glycosides of *Silphium perfoliatum*. II. Chem. Nat. Compd. 20, 636.

Davidyants, E.S., Putieva, Zh.M., Bandyukova, V.A., Abubakirov, N.K., 1984c. Triterpene glycosides of *Silphium perfoliatum*. III. Structure of Silphioside E. Chem. Nat. Compd. 20, 708-710.

Davidyants, E.S., Putieva, Zh.M., Shashkov, A.S., Bandyukova, V.A., Abubakirov, N.K., 1985. Triterpene glycosides of *Silphium perfoliatum*. IV. Structure of Silphioside C. Chem. Nat. Compd. 21, 486-489.

Davidyants, E.S., Putieva, Zh.M., Bandyukova, V.A., Abubakirov, N.K., 1986. Triterpene glycosides of *Silphium perfoliatum*. V. Structure of Silphioside A. Chem. Nat. Compd. 22, 58-60.

De Tommasi, N., Piacente, S., Gacs-Baitz, E., De Simone, F., Pizza, C., Aquino, R., 1998. Triterpenoid saponins from *Spergularia ramosa*. J. Nat. Prod. 61, 323-327.

Duffey, S.S., Stout, M.J., 1996. Antinutritive and toxic components of plant defense against insects. Arch. Biochem. Physiol. 32, 3-37.

El-Sayed, N.H., Wojcinska, M., Drost-Karbowska, K., Williams, J., Mabry, T.J., 2002. Kaempferol triosides from *Silphium perfoliatum*. Phytochemistry 60, 835-838.

Emerenciano V.P., Kaplan, M.A.C., Gottlieb, O.R., 1985. The evolution of sesquiterpene lactones in angiosperms. Biochem. Syst. Ecol. 13, 145-166.

Emerenciano, V.P. Militão, J.S.L.T. Campos, C.C., Romoff, P., Kaplan, M.A.C., Zambon, M., Brant, A.J.C., 2001. Flavonoids as chemotaxonomic markers for Asteraceae. Biochem. Syst. Ecol. 29, 947-957.

Emerenciano, V.P., Bonfanti, M.R.M., Ferreira, Z.S., Kaplan, M.A.C., Gottlieb, O.R., 1986. The evolution of sesquiterpene lactones in Asteraceae. Biochem. Syst. Ecol. 14, 585-589.

Emerenciano, V.P., Cabrol-Bass, M.D., Ferreira, J.P., Alvarenga, S.A.V., Brant, A.J.C., Scotti, M.T., Barbosa, K.O., 2006. Chemical evolution in the Asteraceae. The oxidation–reduction mechanism and production of secondary metabolites. Nat. Prod. Comm. 1, 495-507.

Emerenciano, V.P., Ferreira, Z.S., Kaplan, M.A.C., Gottlieb, O.R., 1987. A chemosystematic analysis of tribes of Asteraceae involving sesquiterpene lactones and flavonoids. Phytochemistry 26, 3103-3115.

Farnsworth, N.R., 1993. Ethnopharmacology and future drug development: the North American experience. J. Ethnopharmacol. 38, 137-143.

Felsenstein, J., 1985. Confidence limits on phylogenetics: an approach using the bootstrap. Evolution 39, 783-791.

Felter, H.W., Lloyd, J.U., 1898. King's American Dispensatory. Henriette's Herbal Homepage: [http://www.henriettesherbal.com/]

Fisher, T.R., Cruden, R.C., 1962. Chromosome numbers and observations in the genus *Silphium*. Ohio J. Sci. 62, 258-259.

Funk, V.A., Bayer, R.J., Sterling, K. Chan, R., Watson, L., Gemeinholder, B., Schilling, E., Panero, J.L., Baldwin, B.G., Garcia-Jacas, N., Susanna, A., Jansen, R.K., 2005. Everywhere but Antarctica: Using a supertree to understand the diversity and distribution of the Compositae. Biol. Skr. 55, 343-374.

Gary, N.E., 1992. Activities and behavior of honey bees. Edited by Graham, J. M. The hive and the honey bee. National Beekeepers Association of New Zealand, Inc.

Gastmans, J.P., Furlan, M., Lopes, M.N., Borges, J.H.G., Emerenciano, V.P., 1990a. A inteligência artificial aplicada à química de produtos naturais. O Programa SISTEMAT. Parte I – Bases teóricas. Quím. Nova 13, 75-80.

Gastmans, J.P., Furlan, M., Lopes, M.N., Borges, J.H.G., Emerenciano, V.P., 1990b. A inteligência artificial aplicada à química de produtos naturais. O Programa SISTEMAT. Parte II – Organização do Programa e Aplicativos. Quím. Nova 13, 75-80.

Gershenzon, J., 1994. Metabolic costs of terpenoid accumulation in higher plants. J. Chem. Ecol. 20, 1281-1328.

Ghisalberti, E., 2004. The Goodeniaceae. Fitoterapia 5, 429-446.

Gonzalez, A.G., Estevez-Reyes, R., Estevez-Braun, A., Ravelo A.G., Jiménez, I.A., Bazzocchi, I.L., Aguilar, M.A., Moujir, L., 1997. Biological activities of some *Argyranthemum* species. Phytochemistry 45, 963-967.

Gottlieb, O.R., 1989. The role of oxygen in phytochemical evolution, towards diversity. Phytochemistry 28, 2545-2558.

Gottlieb, O.R., 1993. Phytochemical evolution: the redox theory. Nat. Prod. Lett. 2, 171-176.

Gottlieb, O.R., 1982. Micromolecular Evolution, Systematics and Ecology. An essay into a novel botanical discipline. Springer-Verlag, Heidelberg.

Grieve, M., 1971 (1931). A Modern Herbal: the medicinal, culinary, cosmetic and economic properties, cultivation and folk-lore of herbs, grasses, fungi, shrubs and trees with their modern scientific uses. Edited by C. F. Leyel. Dover Publications, Inc. New York, New York.

Guo, M., Song, F., Bai, Y., Liu, Z., Liu, S., 2002. Rapid analysis of a triterpenoid saponin mixture from plant extracts by electrospray ionization multi-stage tandem mass spectrometry (ESI-MS^n). Anal. Sci. 18, 481-484.

Hamel, P.B., Chiltoskey, M.U., 1975. Cherokee Plants and Their Uses–a 400 Year History. Herald Publishing Company. North Carolina. 1-65.

Han, L.K., Zheng, Y.N., Yoshikawa, M., Okuda, H., Kimura, Y., 2005. Anti-obesity effects of chikususaponins isolated from *Panax japonicus* rhizomes. BMC Complemen. Altern. Med. 5:9.

Harborne, J.B., 1977. Flavonoids and the evolution of angiosperms. Biochem. Syst. Ecol. 5, 7-22.

Hattori, M., Kawata, Y., Kakiuchi, N., Matsuura, K., Tomimori, T., Namba, T., 1988. Application of liquid chromatography/mass spectrometry to the qualitative analysis of saponins. II. Chem. Pharm. Bull. 36, 4467-4473.

Hegnauer, R., 1996. Phytochemistry and plant taxonomy- an essay on the chemotaxonomic of higher plants. Phytochemistry 25, 1519-1535.

Henry, M., 2005. Saponins and phylogeny: Example of the “gypsogenin group” saponins. Phytochem. Rev. 4, 89-94.

Hilty, J., 2007. “Prairie wildflowers of Illinois”. Illinois wildflowers. http://www.illinoiswildflowers.info/prairie/plant_index.htm

Huhman, D.V., Sumner, L.W., 2002. Metabolic profiling of saponins in *Medicago sativa* and *Medicago truncatula* using HPLC coupled to an electrospray ion-trap mass spectrometer. Phytochemistry 59, 347-360.

Hymete, A., Iversen, T.H., Rohloff, J., Erko, B., 2005. Screening of *Echinops ellenbeckii* and *Echinops longisetus* for biological activities and chemical constituents. Phytomedicine 12, 675-679.

Iturbe-Ormaetxe, I., Harmalampidis, K., Papadopoulou, K., Osbourn, A.E., 2003. Molecular cloning and characterization of triterpene synthases from *Medicago truncatula* and *Lotus japonicus.* Plant Mol. Biol. 51, 731-743.

Jansen, R.K., Palmer, J.D., 1987. A chloroplast DNA inversion marks an ancient evolutionary split in the sunflower family (Asteraceae). Proc. Nat. Acad. Sci. U.S.A. 84, 5818-5822.

Jansen, R.K., Kim K., 1996. Implications of chloroplast DNA data for the classification and phylogeny of the Asteraceae. In D.J.N. Hind and H. Beentje (eds.). Compositae Systematics. Proceedings of the International Compositae Conference, Hind DJN (Editorin-Chief), Vol. 1, Royal Botanic Garden, Kew: 317-339.

Jeffrey, C., 2007. Compositae: Introduction with key to tribes. in: Kadereit, J. W.; Jeffrey, C. (eds.). *Flowering Plants. Eudicots: Asterales. (The families and genera of vascular plants: 8.)* Springer-Verlag, Berlin, Heidelberg.

Jordon-Thaden, I.E., Louda, S.M., 2003. Chemistry of *Cirsium* and *Carduus*: a role in ecological risk assessment for biological control of weeds. Biochem. Syst. Ecol. 31, 1353-1396.

Kapusta, I., Janda, B., Stochmal, A., Oleszek, W., 2005. Determination of saponins in aerial parts of Barrel Medic (*Medicago truncatula*) by liquid chromatography-electrospray ionization/mass spectrometry. J. Agric. Food Chem. 53, 7654-7660.

Karis, P.O., Källersjö, M., Bremer, K., 1992. Phylogenetic analysis of the Cichorioideae (Asteraceae), with emphasis on the Mutisieae. Ann. Mo. Bot. Gard. 79, 416-427.

Keil. D.J., Luckow, M.A., Pinkava, D.J., 1988. Chromosome studies in the Asteraceae from the United States, Mexico, the West Indies, and South America. Am. J. Bot. 75, 652-668.

Kharbanda, R.K., Wallace, S., Walton, B., Donald, A., Cross, J.M., Deanfield, J., 2005. Systemic acyl-coA: cholesterol acyltransferase inhibition reduces inflammation and improves vascular function in hypercholesterolemia. Circulation 111, 804-807.

Kim, K., Jansen, R.K., 1995. ndhF sequence evolution and the major clades of the sunflower family. Proc. Nat. Acad. Sci. U.S.A. 92, 10379-10383.

Kindscher, K., Manfredi, K.P., Britton, M., Demidova, M., Hurlburt, D.P., 1998. Testing prairie plants with ethnobotanical importance for anti-cancer and anti-aids compounds. J. Ethnobiol. 18, 229-245.

Kirk, H., Choi, Y.H., Kim, H.K., Verpoorte, R., van der Meijden, E., 2005. Comparing Metabolomes: the Chemical Consequence of Hybridiztion in Plants. New Phytol. 167, 613-622.

Klinkenborg, V. 2007. "Splendor of the Grass: The Prairie's Grip is Unbroken in the Flint Hills of Kansas", National Geographic feature at: http://ngm.nationalgeographic.com/ngm/0704/feature5/

Kocor, M., St. Pyrek, J., 1973. Triterpenes of *Datura innoxia*. Structure of daturadiol and daturaolone. J. Org. Chem. 38, 3685-3688.

Kowalski R., Wolski, T., 2003. TLC and HPLC analysis of the phenolic acids in *Silphium perfoliatum* L. leaves, inflorescences and rhizomes. J. Planar Chromatogr. Mod. TLC 16, 230-236.

Kowalski, R., 2004. Secondary metabolites in *Silphium integrifolium* in the first 2 years of cultivation. N. Z. J. Crop Hortic. Sci. 32, 397-406.

Kowalski, R., 2007. Studies of selected plant raw materials as alternative sources of triterpenes of oleanolic and ursolic acid types. J. Agric. Food Chem. 55, 656-662.

Kowalski, R., Wiercinski, J., Mardarowicz, M., 2005. Essential oil in leaves and inflorescences of *Silphium integrifolium* Michx. J. Essent. Oil Res.17, 220-222.

Kowalski, R., Wolski, T., 2005. The chemical composition of essential oils of *Silphium perfoliatum* L. Flavor Fragrance J. 20, 306-310.

Latha R.M., Geetha T., Varalakshmi P., 1998. Effect of *Vernonia cinerea* Less flower extract in adjuvant-induced arthritis. Gen pharmacol. 31, 601-6.

Lee, Y.J., Jin, Y.R., Lim, W.C., Ji, S.M., Choi, S., Jang, S., Lee, S.K., 2003. A ginsenoside-Rh1, a component of ginseng saponin, activates estrogen receptor in human breast carcinoma MCF-7 cells. J. Steroid Biochem. Mol. Biol. 84, 463-468.

Levy, M., 1997. Minimum biosynthetic-step indices as measures of comparative flavonoid affinity. Syst. Bot. 2, 89-98.

Li, J., Li, P., Li, H., Song, Y., Bi, Z., Li, Y., 2007. Simultaneous qualification and quantification of eight triterpenoids in *Radix Achyranthis* Bidentatae by high-performance liquid chromatography with evaporative light scattering detection and mass spectrometric detection. J. Sep. Sci. 30, 843-850.

Lin, T., Kondo, N., Shoji, J. 1976. Studies on the constituents of Panacis Japonici Rhizoma. V. The structures of Chikusetsusaponin I, Ia, Ib, IVa and Glycoside P_1. Chem. Pharm. Bull. 24, 253-261.

Linnaeus, C., 1753. *Species Plantarum*. Stockholm: Impensis Laurentii Salvii.

Mabry, T.J., 1973. Is the Centrospermae monophyletic? A review of phylogenetically significant molecular, ultrastructural and other data for Centrospermae families. Pp. 75--85. in: Bednz, G., Santesson, J. & Runnstrom-Reio, V. (eds.). *Chemistry in Botanical Classification*, Nobel Foundation, Stockholm, Sweden and Academic Press, New York.

Mabry, T.J., Bohlmann, F., 1977. Summary of the chemistry of the Compositae. Pp. 1097--1104. in: Heywood, V. Harborne, J. B. & Turner, B. L. (eds.), *The Biology and Chemistry of the Compositae*, Vol. II. Academic Press, London.

Macleod, J.K., Rasmussen, H.B., 1999. A hydroxy-beta-caryophyllene from *Pterocaulon serrulatum*. Phytochemistry 50, 105-108.

Madl, T., Sterk, H., Mittelbach, M., Rechberger. G. N., 2006. Tandem mass spectrometric analysis of a complex triterpene saponin mixture of *Chenopodium quinoa*. J. Am. Soc. Mass Spec. 17, 795-806.

Mahato, S.B., Kundu, A.P., 1994.^{13}C NMR spectra of pentacyclic triterpenoids - a compilation and some salient features. Phytochemistry 37, 1517-1575.

Maisch, J.M., 1881. American Journal of Pharmacy. Vol. 53. Henriette's Herbal Homepage: [http://www.henriettesherbal.com/]

Martinet, A., Ndjoko, K., Terreaux, C., Marston, A., Hostettmann, K., Schutz, Y., 2001. NMR and LC-MS^n characterization of two minor saponins from *Ilex paraguariensis*. Phytochem Anal. 12, 48-52.

Medley, M.E., 1989. *Silphium wasiotensis* (Asteraceae), a new species from the Appalachian plateau in eastern Kentucky. Sida 13, 285-291.

Mehta, B.K., Mehta, D., Itoriya, A., 2004. Structure elucidation by NMR spectroscopy of a new acetylated saponin from *Centratherum anthelminticum*. Carbohydr. Res. 339, 2871-2874.

Mimaki, Y., Harada, H., Sakuma, C., Haraguchi, M., Yui, S., Kudo, T., Yamazaki, M., Sashida, Y., 2004. Contortisiliosides A-G: Isolation of seven new triterpene bidesmosides from *Enterolabium contortisiliquum* and their cytotoxic activity. Helv. Chim. Act. 87, 851-865.

Moerman, D.E., 1991. The medicinal flora of native North America: An analysis. J. Ethnopharmacol. 31, 1-42.

Moerman, D.E., 2003. "Native American Ethnobotany". A database of food, drugs, dyes, fibers of Native American peoples, derived from plants. The University of Michigan-Dearborn. http://herb.umd.umich.edu/

Mohammad, S., Naghmeh, H., Mohammad, K., 2004. Analgesic and anti-inflammatory activity of *Lactuca sativa* seed extract in rats. J. Ethnopharmacol. 92, 325-329.

Mokdad, A.H., Marks, J.S., Stroup, D.F., Gerberding, J.L., 2004. Actual causes of death in the United States, 2000. JAMA. 291,1238-1245.

Morris, L.I., 1881. American Journal of Pharmacy-Vol. 53; *Silphium Laciniatum*, Lin., Rosin Weed. The Southwest School of Botanical Medicine: [http://www.swsbm.com]. 7-11.

Mujoo, K, Haridas, V., Hoffmann, J.J., Wächter, G.A., Hutter, L.K., Lu, Y., Blake, M.E., Jayatilake, G.S., Bailey, D., Mills, G.B., Gutterman, J.U., 2001. Triterpenoid saponins from *Acacia victoriae* (Bentham) decrease tumor cell proliferation and induce apoptosis. Cancer Res. 61, 5486-5490.

Neukirch, H., D'Ambrosio, M., Sosa, S., Altinier, G., Loggia, R.D., Guerriero, A., 2005. Improved anti-inflammatory activity of three new terpenoids derived, by systematic chemical modification from the abundant triterpenes of the flowery plant *Calendula officinalis*. Chem. Biodiversity 2, 657-671.

Nishimura, K., Miyase, T., Noguchi, H., 1999. Triterpenoid saponins from *Ilex kudincha*. J. Nat. Prod. 62, 1128-1133.

Ohana, P., Delmer, D.P., Carlson, R.W., Glushka, J., Azadi, P., Bacic, T., Benziman, M., 1998a. Identification of a novel triterpenoid saponin from *Pisum sativum* as a specific inhibitor of the diguianylate cyclase of *Acetobacter xylinum*. Plant Cell Physiol. 39, 144-152.

Ohana, P., Delmer, D.P., Volman, G., Benziman, M., 1998b. Glycosylated triterpenoid saponin: a specific inhibitor of diguanylate cyclase from *Acetobacter xylinum.* Biological activity and distribution. Plant Cell Physiol. 39, 153-159.

Orians, C.M., 2000. The effects of hybridization in plants on secondary chemistry: implications for the ecology and evolution of plant–herbivore interactions. Am. J. Bot. 87, 1749-1756.

Panero, J.L., 2007. Compositae: Key to the tribes of the Heliantheae alliance. in: Kadereit, J. W.; Jeffrey, C. (eds.). *Flowering Plants. Eudicots: Asterales. (The families and genera of vascular plants: 8.)* Springer-Verlag, Berlin, Heidelberg.

Panero, J.L., Funk, V.A., 2002. Toward a phylogenetic subfamilial classification for the Compositae (Asteraceae). Proc. Biol. Soc. Wash. 115*,* 909-922.

Papadopoulou, K., Melton, R.E., Leggett, M., Daniels, M.J., Osbourn, A.E., 1999. Compromised disease resistance in saponin-deficient plants. Proc. Nat. Acad. Sci. U.S.A. 96, 12939-12928.

Paquette, L.A., Leone-Bay, A., 1983. Triquinane Sesquiterpenes. An iterative, highly stereocontrolled synthese of (±)-silphinene. J. Am. Chem. Soc. 105, 7352-7358.

Parfitt, B.D., 1981. Chromosome number reports LXXXI. Taxon 30, 515-516.

Pari, K., Rao, P.J., Subrahmanyam, B., Rasthogi, J.N., Devakumar, C., 1998. Benzofurans and other constituents of the essential oil of *Ageratum conyzoides*. Phytochemistry 49, 1385-1388.

Passreiter, C.M., 1992. Co-occurrence of 2-pyrrolidineacetic acid with the pyrrolizidines tussilaginic acid and isotussilaginic acid and their 1-epimers in *Arnica* species and *Tussilago farara*. Phytochemistry 31, 4135-4137.

Passreiter, C.M., 1998. Pyrrolizidine alkaloids from *Neurolaena lobata*. Biochem. Syst. Ecol. 26, 839-843.

Pasteels, J.M., Termonia, A., Windsor, D.M., Witte, L., Theuring, C., Hartmann, T., 2001. Pyrrolizidine alkaloids and pentacyclic triterpene saponins in the defensive secretions of *Platyphora* leaf beatles. Chemoecology 11, 113-120.

Pcolinski, M.J., Doskotch, R.W., Lee, A.Y., Clardy, J., 1994. Chlorosilphanol A and silphanepoxol, labdane diterpenes from *Silphium perfoliatum*. J. Nat. Prod. 57, 776-783.

Perret, C., Wolfender, J.L., Hostettmann, K., 1999. LC/ES-MS analysis of triterpene glycosides: rapid estimation of the saponin content of dried berries of *Phytolacca dodecdandra.* Phytochem. Anal. 10, 272-278.

Perry, L.M., 1937. Notes on *Silphium*. Rhodora 39, 281-297.

Petrovic, S.D., Gorunovic, M.S., Wray, V., Merfort, I., 1999. A taraxasterol derivative and phenolic compounds from *Hieracium gymnocephalum*. Phytochemistry 50, 293-6.

Plasman, V., Plehiers, M., Braekman, J.C., Daloze, D., de Biseau, J.C., Pasteels, J.M., 2001. Chemical defenses in *Playtphora kollari* Baly and *Leptinotarsa behrensi* Harold (Coleoptra: Chrosmelidae). Hypothesis on the origin and evolution of leaf beetle toxins. Chemoecology 11, 107-112.

Pleasants, J.M., Jurik, T.W., 1992. Dispersion of seedlings of the prairie compass plant, *Silphium laciniatum* (Asteraceae). Am. J. Bot. 79, 133-137.

Porcher, F. P., 1869. Resources of the Southern Fields and Forests, Medical, Economical and Agricultural; being also a Medical Botany of the Southern States; with Practical Information on the Useful Properties of the Trees, Plants, and Shrubs. Walker, Evans & Cogswell, Printers. 460-461.

Proksch, P., Rodgriguez, E.1983. Chromenes and benzofurans of the Asteraceae, their chemistry and biological signficance. Phytochemistry 11, 2335-2348.

Quetin-Leclercq, J., Elias, R., Balansard, G., Bassleer, R., Angenot, L., 1992. Cytotoxic activity of some triterpenoid saponins. Planta Med. 58, 279-281.

Rao, A.V., Gurfinkel, D.M., 2000. The bioactivity of saponins: triterpenoid and steroidal glycosides. Drug Metab. Drug Interact. 17, 211-235.

Rhaman, A., Ahamed, A., Amakawa, T. Goto, N. and Tsurumi, S., 2001. Chromosaponin I specifically interacts with AUX1 protein in regulating the gravoptropic response of *Arabidopsis roots.* Plant Physiol. 125, 990-1000.

Richardson, P.M., Young, D.A., 1982. The phylogenetic content of flavonoid point scores. Biochem. Syst. and Ecol. 10, 251-255.

Rickett, H.W., 1967. Wild Flowers of the United States: Southeastern States. The New York Botanical Garden/McGraw-Hill, New York, p. 580.

Rieseburg, L.H., Michelmore, R. 2003. Compositae Genome Project. [http://compgenomics.ucdavis.edu/index.htm].

Rivier, L., Bruhn, J.G., 1982. J. Ethnopharmacol. 5, 303-358.

Rao. A.V., Gurfinkel, D.M., 2000. The bioactivity of saponins: triterpenoid and steroidal glycosides. Drug Metab. Drug Interact. 17, 211-235.

Roeder, E., Wiedenfeld, H., Hille, T., Britz-Kirstgen, R., 1984. Pyrrolizidine alkaloids in *Echinacea angustifolia* DC. Und *Echinaea purperea* M. Dtsch. Apoth. Ztg. 124, 2316-2318.

Samson, F.B., Knopf, F.L., 1994. Prairie conservation in North America. BioScience 44, 418-421.

Santos Rosa, C., Garcia Gimenez, M.D., Saenz Rodriguez, M.T., de la Puerta Vazquez, R., 2007. Antihistaminic and antieicosanoid effects of oleanolic and ursolic acid fraction from *Helichrysum picardii*. Pharmazie 62, 459-462.

Sarker, S.D., Laird, A., Nahar, L., Kumarasamy, Y., Jaspars, M., 2001. Indole alkaloids from the seeds of *Centaurea cyanus* (Asteraceae). Phytochemistry 57, 1273-1276.

Sayre, L.E., 1917. A Manual of Organic Materia Medica and Pharmacognosy. Henriette's Herbal Homepage: [http://www.henriettesherbal.com/]

Schroeder, P., Luckner, M., 1968. Physiology of formation of the quinoline alkaloid echinorine in *Echinops ritro*. Planta Medica 16, 99-108.

Schüngel, J., Passreiter, C.M., 2000. 2-pyrrolidineacetic acid and pyrrolizidine alkaloids from *Melampodium divaricatum*. Biochem. Syst. Ecol. 28, 705-706.

Scudder, J.M., 1870. Specific Medication and Specific Medicines. Henriette's Herbal Homepage: [http://www.henriettesherbal.com/]

Seaman, F.C., Funk, V.A., 1983. Cladistic - Analysis of complex natural developing transformation series from sesquiterpene lactone data. Taxon 32, 1-27.

Seamen, F., Bohlmann, F., Zdero, C., Mabry, T.J. 1990. *Diterpenes of flowering plants: Compositae (Asteraceae)*. Springer-Verlag, New York.

Seigler, D.S., 1998. *Plant Secondary Metabolism*. Kluwer Academic, Boston.

Settle, W.J., 1967. The chromosome morphology in the genus *Silphium*. Ohio J. Sci. 67, 10-19.

Settle, W.J., Fisher, T.R., 1972. A biosystematic study of *Silphium integrifolium* Michaux (Compositae). Ohio J. Sci. 72, 254-264.

Shemluck, M., 1982. Medicinal and other uses of the Compositae by Indians in the United States and Canada. J. Ethnopharmacol. 5, 303-358.

Shimizu, S., Ishihara, N., Umehara, K., Miyase, T., Ueno, A., 1988. Sesquiterpene glycosides and saponins from *Cynara cardunculus* L. Chem. Pharm. Bull. 36, 2466-2474.

Sigstad, E.E., Cuenca, M.D., Catalan, C.A.N., Gedris, T.E., Herz, W., 1999. Clerodanes from *Onoseris alata*. Phytochemistry 50, 835-838.

Small, J.K., 1933. *Silphium*. Manual of the southeastern flora. University of North Carolina Press, Chapel Hill, NC. 1408-1415.

Smith, E.B., Hyatt, P.E., Golden, K.D., 1992. Documented chromosome numbers 1002: I chromosome numbers of some Arkansas flowering plants. Sidas 15, 145-146.

Sneath, P.H.A., Sokal, R.R., 1973. *Numerical Taxonomy*. WH Freeman, San Francisco.

Stuessy, T.F., Crawford, D.J., 1983. Flavonoids and phylogenetic reconstruction. Plant Syst. Evol. 143, 83-107.

Swofford, D.L., 2002. PAUP*: phylogenetic analysis using parsimony (*and other methods), version 4. Sinauer Associates, Sunderland, Massachusetts, USA.

Syrov, V.N., Khushbaktova, Z.A., Davidyants, E.S., 1992. Triterpenoid glycosides from *Silphium perfoliatum* L. Hypolipidemic activity of silphioside. Khim. Farm. Zh. 26, 66-9.

Szakiel, A., Wasiukiewicz, I., Janiszowska, W., 1995. Metabolism of [3-^{3}H] oleanolic acid in the isolated *Calendula officinalis* leaf cells and transport of the synthesized glycosides to the cell wall and the extracellular space. Acta Biochem. Polon. 42, 25-29.

Szakiel, A., Ruszkowski, D., Janiszowska, W., 2003. Excretion of oleanolic acid glycosidesto the medium from roots of marigold *Calendula officinalis* L. Pol. J. Nat. Sci. 1, Suppl. 1, 216.

Szakiel, A., Ruszkowski, D., Janiszowska, W., 2005. Saponins in *Calendula officinalis* L.- structure, biosynthesis, transport and biological activity. Phytochem. Rev. 4, 151-158.

Taylor, W.R., 1926. Chromosome morphology in *Frittlaria*, *Alstroemeria*, *Silphium*, and other genera. Am. J. Bot. 13, 179-193.

Teeri, T.H., Elomas, P., Kotilainen, M., Albert, V.A., 2006. Mining plant diversity: *Gerbera* as a model system for plant developmental and biosynthetic research. BioEssays 28, 756-767.

Teixeira da Silva, J.A., 2004. Mining the essential oils of the Anthemideae. Afr. J. Biotechnol. 3, 706-720.

Teixeira da Silva, J.A., Yonekura, L., Kaganda, J., Mookdasanit, J., Nhut, D.T., Afach, G., 2004. Important secondary metabolites and essential oils of species within the Anthemideae (Asteraceae). J. Herb. Spice. Medic. Plants 11, 1-45.

Tooker, J.F., Koenig, W.A., Hanks, L.M., 2002. Altered host plant volatiles are proxies for sex phermones in the gall wasp *Antistrophus rufus*. Proc. Nat. Acad. Sci. U.S.A. 99, 15486-15491.

Tooker, J.F., Hanks, L.M., 2004. Endophytic insect communities of two prairie perennials (Asteraceae: *Silphium* spp.). Biodiversity and Conserv. 13, 2551-2566.

Tooker, J.F., Hanks, L.M., 2006. Tritrophic interactions and reproductive fitness of the prairie perennial *Silphium laciniatum* Gillette (Asteraceae). Environ. Entom. 35, 537-545.

Tooker, J.F., Crumrin, A.L., Hanks, L.M., 2005. Plant volatiles are behavioral cues for adult females of the gall wasp *Antistrophus rufus*. Chemoecology 15, 85-88.

Ukiya, M., Akihisa, T., Tokuda, H., Suzuki, H., Mukainaka, T., Ichiishi, E., Yasukawa, K., Kasahara, Y., Nishino, H., 2002. Constituents of Compositae plants. 3. Anti-tumor promoting effects and cytotoxic activity against human cancer cell lines of triterpene diols and triols from edible *Chrysanthemum* flowers. Cancer Lett. 177, 7-12.

Ukiya, M., Akihisa, T., Yasukawa, K., Kasahara, Y., Kimura, Y., Koike, K, Kikaido, T., Takido, M., 2001. Constituents of Compositae plants. 2. Triterpene diols, triols, and their 3-O-fatty acid esters from edible *Chrysanthemum* flower extract and their anti-inflammatory effects. J. Agric. Food Chem. 49, 3187-3197.

Valant-Vetschera, K.M., Wollenweber, E., 2007. Exudate flavonoids in miscellaneous Asteraceae. Chemodiversity of exudate flavonoids in seven tribes of Cichorioideae and Asteroideae (Asteraceae). Journal of Biosci. 62, 155-163.

Van Wyk, B.E., 2003. The value of chemosystematics in clarifying relationships in the genistoid tribes of the papilionoid legumes. Biochem. Syst. Ecol. 31, 875-884.

Vichnewski, W., Skrochy, C.A., Maria, A., Nasi, T.T., Lopes, J.L.C., Herz, W., 1999. 15-hydroxyeremantholide B and derivatives from *Eremanthus arboreus*. Phytochemistry, 50, 317-320.

Vidal-Ollivier, E., Balansard, G., 1989. Revised structures of triterpenoid saponins from the flowers of *Calendula officinalis*. J. Nat. Prod. 52, 1156-1159.

Vincken, J., Heng, L., de Groot, A., Gruppen, H., 2007. Saponins, classification and occurrence in the plant kingdom. Phytochemistry 68, 275-297.

Wagenitz, G., 1976. Systematics and phylogeny of the Compositae (Asteraceae). Plant Syst. Evol. 125, 29-46.

Wang, F., Hua, H., Pei, Y., Chen, D., Jing, Y., 2006. Triterpenoids from the resin of *Styrax tonkinensis* and their anti-proliferative and differentiation effects in human leukaemia HL-60 cells. J. Nat. Prod. 69, 807-810.

Wang, K., Mao, J., Tai, Y., Pan, Y., 2006. Novel skeleton terpenes from *Celastrus hypoleucus* with anti-tumor activities. Bioorg. Med. Chem. Lett. 16, 2274-2277.

Waterman, P. G. & Gray, A. I. 1987. Chemical systematics. Nat. Prod. Rep. 4, 175-203.

Weaver, J. E., 1954. North American Prairie. Johnsen Publishing, Lincoln, NE.

Williams, J. D., 2006. The flavonoids and phenolic acids of the genus *Silphium* and their chemosystematic and medicinal value. Texas, USA: University of Texas at Austin (PhD. dissertation).

Winston, M. L., 1987. The biology of the honey bee. Harvard University Press.

Wojcinska, M., Drost-Karbowska, K., 1998. Phenolic acids in *Silphium perfoliatum* L. flowers (Asteraceae: Compositae). Pol. Pharm. Soc. 55, 413-416.

Wojcinska, M., Williams, J., Mabry, T.J., Ahmed, A.A., Davis, B.D., Toth, G., El-Sayed, N.H., Matlawska, I., Clevinger, J., 2006. Flavonol triglycosides from the leaves of *Silphium albiflorum*. Nat. Prod. Comm. 1, 941-948.

Wollenweber, E., Roitman, J.N., 2007. New reports on surface flavonoids from *Chamaebatiaria* (Rosaceae), *Dodonaea* (Sapindaceae), *Elsholtzia* (Lamiaceae), and *Silphium* (Asteraceae). Nat. Prod. Comm. 2, 385-389.

Wolski, T., Kowalski, R., Mardarowicz, M., 2000. Chromatographic analysis of essential oil occurring in inflorescences, leaves and rhizomes of *Silphium perfoliatum* L. Herba Pol. 46, 235-242.

Xiang, T., Shibuya, M., Katsube, Y., Tsutsumi, T., Otsuka, M., Zhang, H., Masuda, K., Ebizuka, Y. 2006. A new triterpene synthase from *Arabidopsis thaliana* produces a tricylic triterpene with two hydroxyl groups. Org. Lett. 8, 2835-2838.

Yoshikawa, M., Murakami, T., Kishi, A., Kageura, T., Matsuda, H., 2001. Medicinal Flowers. III. Marigold. : Hypoglycemic, gastric empyting inhibitory, and gastroprotective principles and new oleanane-type triterpene oligoglycosdies, calendasaponins A, B. C and D, from Egyptian *Calendula officinalis*. Chem. Pharm. Bull. 49, 863-870.

Zdero, C., Bohlmann, F., 1990. Systematics and evolution within the Compositae, seen with the eyes of a chemist. Plant Syst. Evol. 171, 1-14.

VITA

Lalita Maria Calabria was born in Perkasie, Pennsylvania on May 22, 1980, the daughter of Mary Catherine Calabria and Charles Angelo Calabria. After completing her education at Archbishop Wood High School in 1998, she entered The Evergreen State College in Olympia, Washington, where she received her Bachelor of Science, double majoring in Botany and Plant Ecology (June 2002). During the following year and a half she was employed as a field botanist at the Mount Baker-Snoqualmie National Forest in Washington and later as a natural product chemist and lichen taxonomist/curator at Oregon State University in Corvallis, Oregon. In September 2003, she began Graduate School at the University of Texas at Austin. Since entering the Plant Biology Graduate Program, Lalita has been a teaching assistant for several botany and biology related courses. In addition, she has served as a research assistant in Dr Mabry's phytochemical laboratories and has supervised dozens of students who obtained credit working as undergraduate research assistants in Dr Mabry's lab. Her doctoral studies were supported by awards including, a University Continuing Fellowship, several departmental and program travel funds and the Welch Foundation (F-130) grant to Dr. Tom J. Mabry.

Permanent Address: 511 E. 40th St., Austin, Texas 78751

This dissertation was typed by the author.